

AI-Assisted COPD Diagnosis Under the GOLD 2026 Framework: A Narrative Review of Opportunities, Evidence Gaps, and Implementation Pathways

Phalguni Deswal*

Independent Researcher, India.
*Corresponding Author Email: phalgunideswal@gmail.com

(Received: 12.10.2026; Accepted: 20.11.2026)

Abstract

Chronic obstructive pulmonary disease (COPD) is still widely underdiagnosed, misdiagnosed, and diagnosed late, mainly because spirometry is underused and often poorly performed or interpreted. The 2026 report of the Global Initiative for Chronic Obstructive Lung Disease (GOLD) keeps post-bronchodilator spirometry as the diagnostic anchor, but it adds targeted case finding, opportunistic detection through lung cancer screening, a clearer role for computed tomography, the concept of disease activity and, for the first time, a chapter on artificial intelligence (AI) and emerging technologies. This narrative review examines how AI could support COPD diagnosis within that framework and how far the published evidence actually goes. The paper mapped studies across five data streams: spirometry quality control and interpretation; chest computed tomography; chest radiographs and electrocardiograms; physiological and acoustic signals; and electronic health records, including large language models. The most mature evidence concerns spirometry interpretation software and deep learning on computed tomography, where several models show good discrimination and some external validation. Most other approaches remain at the stage of retrospective development in small or selected samples. Common weaknesses across modalities include reference standards that do not match the GOLD definition, label noise from diagnostic codes, spectrum bias, poor reporting of calibration, scarce data from low- and middle-income countries, and an almost complete absence of prospective studies with patient-relevant outcomes. The paper proposes a staged implementation pathway in which AI works as a triage and quality layer that sends the right people to confirmatory spirometry rather than replacing it. This should be supported by local validation, monitoring for dataset shift, equity audits, and AI-specific reporting standards. Pragmatic trials measuring diagnostic yield, time to diagnosis, and downstream care are now the priority.

Keywords: *Chronic obstructive pulmonary disease; GOLD 2026; Artificial intelligence; Spirometry; Case finding; Quantitative computed tomography*

INTRODUCTION

Chronic obstructive pulmonary disease (COPD) affects several hundred million people and is among the leading causes of death worldwide. A pooled analysis of population studies estimated that about 10% of adults aged 30 to 79 years had COPD in 2019, with most of the burden in low- and middle-income countries (Adeloye *et al.*, 2022). The Burden of Obstructive Lung Disease (BOLD) study showed, much earlier, that prevalence based on spirometry is far higher than prevalence based on doctor diagnosis (Buist *et al.*, 2007). That gap has not closed. Roughly 70% of people with COPD worldwide have never been diagnosed, and between 30% and 60% of people who carry the label do not have airflow obstruction when spirometry is finally done (Diab *et al.*, 2018; Agusti *et al.*, 2026). A recent population study in Catalan primary care showed the same pattern of poor diagnostic accuracy in routine practice (Vila *et al.*, 2025).

These failures matter as people with undiagnosed obstructive lung disease report more symptoms, poorer quality of life, more health care use, and lower work productivity than people without disease (Gerstein *et al.*, 2023). Misdiagnosis can lead to prescribing inhalers that are not needed or missing alternative diagnoses such as heart failure. Late diagnosis means that treatment starts when much of the lung damage has already taken hold, usually in the sixth or seventh decade of life (Martinez *et al.*, 2022). The Lancet Commission on COPD argued that the disease should be thought of as a lifelong process with many causes, and that earlier

recognition is central to any serious plan to reduce its burden (Stolz *et al.*, 2022). Lung function in adulthood depends partly on how well the lungs grew in early life, so not every person with low lung function followed the classic path of accelerated decline (Agusti and Faner, 2019; Christenson *et al.*, 2022).

The Global Initiative for Chronic Obstructive Lung Disease (GOLD) has set the reference framework for COPD for more than two decades. Its 2026 report, released in November 2025, does not change the definition of the disease, but it changes the emphasis (Global Initiative for Chronic Obstructive Lung Disease, 2025a; Venkatesan, 2025). It moves from population screening towards targeted case finding, it names lung cancer screening as an opportunity to find unrecognised COPD, it gives computed tomography (CT) a clearer role when findings do not fit spirometry, and it introduces disease activity as something clinicians should look for and treat (Global Initiative for Chronic Obstructive Lung Disease, 2025b; Cazzola *et al.*, 2026). It also includes, for the first time, a full chapter on artificial intelligence (AI) and emerging technologies. Members of the GOLD Science Committee later expanded this chapter into a perspective that sets out where AI might help with underdiagnosis, misdiagnosis, and late diagnosis (Agusti *et al.*, 2026).

AI is not new to respiratory medicine. Machine learning has been applied to pulmonary function tests for years (Das *et al.*, 2018), and broader reviews have described both the promise of deep learning in medicine and the gap between good performance on a test set and real benefit to patients (Topol, 2019; Kelly *et al.*, 2019). What is new is that a major guideline body now treats AI as part of the diagnostic conversation. This raises practical questions. Which AI tools fit the GOLD 2026 pathway, and at which step? How strong is the evidence for each? What would need to be true before a health system could use them safely, especially in settings where spirometry is scarce?

This review tries to answer those questions. It has four aims. The first is to summarise the diagnostic elements of GOLD 2026 that create openings for AI. The second is to map the evidence for AI across the main data streams that could feed COPD diagnosis. The third is to appraise that evidence critically, with attention to reference standards, validation, calibration, equity and outcomes. The fourth is to propose implementation pathways and a research agenda that keep spirometry at the centre while using AI to decide who should be tested and to improve the quality of testing.

APPROACH TO THE REVIEW

This is a narrative review, not a systematic review. A narrative design is chosen because the field is moving fast, the studies use very different data types and designs, and the questions of interest are partly conceptual. To keep the narrative approach disciplined, we followed the items of the Scale for the Assessment of Narrative Review Articles (SANRA), which asks for a clear justification, explicit aims, a described literature search, appropriate referencing, sound scientific reasoning, and relevant presentation of data (Baethge *et al.*, 2019). A similar narrative approach, based on a PubMed search, was used by the GOLD Science Committee for its own AI perspective (Agusti *et al.*, 2026).

PubMed and Google Scholar were searched, and the reference lists of key papers were screened for work published between January 2015 and October 2026. Search terms combined COPD or chronic obstructive pulmonary disease with artificial intelligence, machine learning, deep learning, neural network, large language model, spirometry, computed tomography, chest radiograph, electrocardiogram, voice, respiratory sound, electronic nose, oscillometry, and electronic health record. A review of the GOLD 2026 report, its summary of key changes, and published commentaries on it was carried out. Priority was given to peer-reviewed studies that addressed detection, diagnosis, quality control or interpretation, and to studies with external validation or a comparison with clinicians. Studies on exacerbation prediction and remote monitoring were included only where they shed light on diagnostic methods or on general problems with AI evidence.

To appraise the evidence, the domains of the Prediction Model Risk of Bias Assessment Tool (PROBAST) were used as a lens rather than as a formal scoring tool (Wolff *et al.*, 2019), and studies were checked against the reporting expectations of TRIPOD+AI (Collins *et al.*, 2024). Each line of evidence was grouped into one of four broad levels of maturity: model development with internal validation only; external validation on independent data; prospective evaluation of diagnostic accuracy in the intended setting; and evaluation of clinical impact, which is the stage addressed by DECIDE-AI and CONSORT-AI (Vasey *et al.*, 2022; Liu *et al.*, 2020). These levels are used throughout the paper and in Table 2.

THE GOLD 2026 DIAGNOSTIC FRAMEWORK

Any useful AI tool for COPD has to fit the way the disease is defined and diagnosed. This section sets out the parts of GOLD 2026 that matter for diagnosis and shows where each one creates a task that AI might help with. The main points are summarised in Table 1.

Spirometry Remains the Anchor

GOLD 2026 keeps the established definition. A diagnosis of COPD needs a post-bronchodilator ratio of forced expiratory volume in one second to forced vital capacity (FEV1/FVC) below 0.70 in a person with compatible symptoms and exposures (Global Initiative for Chronic Obstructive Lung Disease, 2025a). What has changed is the pathway towards that test. Following a GOLD Science Committee review, pre-bronchodilator spirometry is now accepted as a first step in symptomatic people. A normal pre-bronchodilator result usually rules out COPD, while an abnormal one must be confirmed after a bronchodilator (Singh *et al.*, 2025). People whose obstruction disappears after the bronchodilator are treated as at risk rather than healthy, because they appear to decline faster over time (Cazzola *et al.*, 2026). Routine reversibility testing to separate asthma from COPD is discouraged.

This change has two consequences for AI. First, more tests may be done in primary care and community settings, where

Table 1: Diagnostic elements of GOLD 2026 and the tasks they create for AI.

GOLD 2026 element	What it asks of clinicians	Main bottleneck in practice	Possible AI contribution	Evidence maturity
Post-bronchodilator FEV1/FVC < 0.70	Confirm persistent obstruction with good-quality spirometry	Poor test quality and interpretation in primary care	Automated quality grading; interpretation support	External validation; one blinded primary care study
Pre-bronchodilator spirometry as first step	Use simpler testing to rule out COPD in symptomatic people	Volume of tests; variable technique	Real-time quality feedback; triage of abnormal results	Development and external validation
Targeted case finding	Identify symptomatic or exposed people for testing	Low sensitivity of simple tools; poor use of records	Risk scores from health records; flags from routine ECG or radiographs	Mostly development; some external validation
Opportunistic detection in lung cancer screening	Use existing CT to find unrecognised COPD	Radiologist time; inconsistent reporting of emphysema	Deep learning on low-dose CT; automated emphysema and airway metrics	External validation; prospective pilots without AI
CT when findings are discordant	Characterise emphysema, airways, mucus plugs	Limited quantitative CT outside research centres	Automated quantification and phenotyping	Cohort-based validation
Pre-COPD and PRISm	Recognise at-risk states without a COPD label	Single time-point testing; no agreed management	Trajectory models; imaging subtypes	Development only
Disease activity	Track progression and instability over time	Sparse longitudinal data	Longitudinal prediction from records and repeated tests	Development only

technical quality is often poor. International standards define how acceptable and repeatable manoeuvres should look (Graham *et al.*, 2019), and how results should be interpreted (Stanojevic *et al.*, 2022). Meeting those standards at scale is hard, which is exactly where automated quality control could add value. Second, interpretation depends on reference equations. The Global Lung Function Initiative equations (Quanjer *et al.*, 2012) were long used with race-specific coefficients. An official American Thoracic Society statement in 2023 recommended moving away from race and ethnicity in pulmonary function interpretation (Bhakta *et al.*, 2023), and race-neutral approaches have since been developed (Bowerman *et al.*, 2023). Any AI tool trained on historical reports will have learned whichever convention was in use, a point on which the discussion will return in the later part of the paper.

From Screening to Targeted Case Finding

GOLD 2026 explicitly advises against spirometric screening of asymptomatic, low-risk adults. It recommends targeted case finding in people with persistent respiratory symptoms or clear risk factors, such as tobacco smoking, biomass smoke, occupational dusts, recurrent infections or early-life lung injury (Global Initiative for Chronic Obstructive Lung Disease, 2025a; Cazzola *et al.*, 2026). The report supports short questionnaires, handheld devices such as peak flow meters and microspirometers, and opportunistic testing in high-yield settings.

The performance of simple tools sets a useful benchmark for AI. In a large United States primary care study of 4,325 patients, the CAPTURE questionnaire combined with peak flow identified only 48% of those with clinically significant undiagnosed COPD, although its specificity was close to 89% (Martinez *et al.*, 2023). The yield of case finding is therefore modest when it relies on a few variables. On the other hand, the UCAP randomised trial showed that finding and treating previously undiagnosed asthma or COPD in symptomatic adults reduced respiratory health care use and improved symptoms compared with usual care (Aaron *et al.*, 2024b). Together with an analysis of the costs and benefits of case finding (Aaron *et al.*, 2024a), this trial gives the strongest argument so far that earlier diagnosis is worth pursuing. AI enters here as a way to make case finding more selective, by using information already stored in records, images, and routine tests.

Imaging, Structural Abnormalities and Early Disease

GOLD 2026 does not recommend routine CT for diagnosis. It does advise CT when symptoms or exacerbations seem out of proportion to spirometry, when incidental abnormalities are found, or when lung volume reduction is being considered (Cazzola *et al.*, 2026). It also highlights lung cancer screening programmes as a place to find patients with COPD who have not been diagnosed (Global Initiative for Chronic Obstructive Lung Disease, 2025a). The idea is not new. More than a decade ago, quantitative emphysema and air-trapping measures on screening CT were shown to identify a large share of participants with airflow limitation (Mets *et al.*, 2011), and incidental emphysema on CT predicts faster decline in lung function and later obstruction (Koo *et al.*, 2016). In young smokers with normal spirometry, structural CT changes already predict lung function decline (Ritchie *et al.*, 2024). A policy perspective has argued that CT done for other reasons should be used more actively to detect COPD and related chronic diseases (Labaki *et al.*, 2024), and a prospective pilot within a lung cancer screening programme has shown that such case finding is feasible (Spetrini *et al.*, 2025).

Pre-COPD and PRISm

Since 2023, GOLD has used the terms pre-COPD and preserved ratio impaired spirometry (PRISm) for people who have symptoms, functional or structural abnormalities, or a low FEV1 with a normal FEV1/FVC ratio (Agustí *et al.*,

2023). The evidence behind these terms comes from cohorts showing that many smokers with normal spirometry have emphysema, airway wall thickening, exacerbations and activity limitation (Regan *et al.*, 2015; Woodruff *et al.*, 2016), and that people with PRISm move between categories over time and have higher mortality (Wijnant *et al.*, 2020). The conceptual case for treating these states as early disease was set out by Han *et al.* (2021). In 2025, a multidimensional diagnostic scheme from the COPDGene and CanCOLD investigators combined symptoms, quality of life, spirometry and CT abnormalities. It classified some people without spirometric obstruction as COPD who nevertheless had higher risks of death, exacerbations, and rapid decline (Bhatt *et al.*, 2025). Previews of the GOLD 2027 report indicate that it will describe six groups that can be categorised as pre-COPD, including people with CT evidence of emphysema or airway wall thickening (Healio, 2026). This growing role for structure and trajectory, rather than a single ratio, is where imaging AI and longitudinal models are most relevant (Melén *et al.*, 2024).

Disease Activity

GOLD 2026 adopts disease activity as a therapeutic target, borrowing a concept from other chronic inflammatory diseases. Activity refers to biological processes that drive progression and that may respond to treatment. The report defines disease stability as a state with no exacerbations, no worsening of symptoms, and no accelerated loss of lung function (Global Initiative for Chronic Obstructive Lung Disease, 2025b). This is not a diagnostic criterion, but it changes what diagnosis is for. A diagnosis is now meant to open the door to early and continued assessment of activity. Commentators have noted that the concept still lacks validated biomarkers beyond blood eosinophils (Cazzola *et al.*, 2026). Detecting activity requires repeated measurements over time, which is a natural task for models that work on longitudinal data.

The New Chapter on Artificial Intelligence

Chapter 6 of GOLD 2026 is the first chapter in the report devoted to AI and emerging technologies. It describes possible roles in reducing underdiagnosis, improving diagnostic precision and refining phenotyping through analysis of health records, spirometry and imaging, together with uses in risk prediction, remote monitoring and self-management (Global Initiative for Chronic Obstructive Lung Disease, 2025a; Cazzola *et al.*, 2026). Among AI-assisted alternatives to spirometry, it lists respiratory sounds and voice, oscillometry, motion sensors, electronic nose signals, photoplethysmography, chest CT and electrical impedance tomography. The chapter stresses that most applications are early, that data are fragmented and poorly interoperable, that external validation is limited, and that bias may harm under-represented groups. It does not endorse any specific product. The companion perspective by the GOLD Science Committee adds a list of key challenges, including limited prospective validation, uncertain clinical benefit, restricted transferability, dependence on high-quality data, poor transparency, and the risk of reinforcing disparities (Agusti *et al.*, 2026). It also states that guidelines remain essential in the era of AI, because they are a structured form of human consensus that AI systems themselves can draw on.

AI ACROSS THE DIAGNOSTIC PATHWAY

AI can enter the COPD pathway at three points: before testing, by deciding who should be tested; during testing, by improving the quality and reading of spirometry; and after testing, by adding structural or longitudinal information that the ratio alone cannot give. The evidence for each data stream is reviewed below, and selected studies are summarised in Table 2.

Spirometry Quality Control and Interpretation

Spirometry is effort-dependent, and its quality in routine primary care is often poor. In the United Kingdom, only a small minority of primary care tests meet all international technical criteria, and agreement between primary care staff and specialists on interpretation is weak (Sunjaya *et al.*, 2025). Because GOLD 2026 makes spirometry both the first step and the confirmatory step, errors at this stage propagate through the entire pathway.

Quality Control

Deep learning applied to flow–volume and volume–time curves can grade manoeuvres against the acceptability and usability criteria of the American Thoracic Society and European Respiratory Society. Das *et al.* (2020) showed that such an algorithm can help standardise these judgements, which are otherwise subjective and vary between technicians. Wang *et al.* (2022c) combined spirometric indices with the curves themselves to automate quality assurance. Related work has used deep learning on flow–volume curves to identify ventilatory patterns (Wang *et al.*, 2022a) and to detect upper airway obstruction (Wang *et al.*, 2022b), which is a common mimic that simple ratio-based rules miss. Newer signal-processing approaches combine adaptive Fourier decomposition with deep learning to improve interpretation (Li *et al.*, 2025). The practical attraction of quality control is that it acts during the test, so a poor manoeuvre can be repeated while the patient is still present.

Interpretation Support

The best-known study in this area compared AI software with 120 pulmonologists from 16 European hospitals, each interpreting the same 50 cases of full pulmonary function tests with clinical information. The pulmonologists matched guideline-based pattern interpretation in about 74% of cases and reached the correct diagnosis in about 45%, with large disagreement between readers. The software matched the patterns in every case and reached the correct diagnosis in 82% (Topalovic *et al.*, 2019). The cases were selected, and the setting was experimental, but the study made clear how variable expert reading can be. A later study showed that pulmonologists became more accurate when working together with an explainable version of the software, which matters because most clinicians will want to see why a suggestion is made (Das *et al.*, 2023).

The step to primary care is harder. Sunjaya *et al.* (2025) tested a spirometry-only version of the software on 1,113 patients from primary care clinics in north-west London, with a

Table 2: Selected studies of AI and comparator tools for COPD detection, ordered by data source.

Study	Data source	Population and setting	Reference standard	Main reported result	Maturity level
Martinez *et al.* (2023), non-AI benchmark	CAPTURE questionnaire and peak flow	4,325 US primary care patients	Spirometry, clinically significant COPD	Sensitivity 48.2%, specificity 88.6%	Prospective accuracy
Topalovic *et al.* (2019)	Full pulmonary function tests with clinical data	50 cases read by 120 pulmonologists, 16 European hospitals	Guideline patterns; expert diagnosis	Diagnosis correct in 82% (AI) vs 45% (pulmonologists)	Retrospective reader comparison
Sunjaya *et al.* (2025)	Primary care spirometry and demographics	1,113 patients, London primary care	Expert panel with full records	High sensitivity and specificity for COPD; weaker for other diseases	External, blinded, retrospective
Cosentino *et al.* (2023)	Raw spirograms	UK Biobank	Spirometry, records, genetics	Improved COPD risk models; predicted survival and exacerbations	Development, large cohort
Tang *et al.* (2020)	Low-dose screening CT	2,589 smokers in lung cancer screening	Spirometry	AUC 0.87 vs 0.68 for mean lung density	Development with held-out testing
Gonzalez *et al.* (2018)	Chest CT	COPDGene and ECLIPSE cohorts	Spirometry; follow-up events	Detected and staged COPD; predicted exacerbations and death	Validation across two cohorts
Schroeder *et al.* (2020)	Chest radiographs	6,749 exams, single US centre	FEV1/FVC < 0.70	AUC 0.81 vs 0.70 for report text	Single-centre retrospective
Ueda *et al.* (2024)	Chest radiographs	141,734 image–spirometry pairs, five Japanese sites	Measured FEV1	Correlation 0.92	Multi-institutional retrospective
Vaid *et al.* (2026)	12-lead ECG	208,231 ECGs; external hospital and UK Biobank	Diagnostic codes	AUC 0.80 internal, 0.82 external, 0.75 UK Biobank	External, retrospective
Idrisoglu *et al.* (2024)	Voice recordings	Purpose-collected dataset	Clinical COPD diagnosis	Automated classification feasible	Development only

reference diagnosis set by expert pulmonologists who had access to full medical records. The software identified COPD with high sensitivity and specificity, but performed less well for other chronic respiratory diseases. It had been trained on secondary care data, which contains fewer normal tests than an unselected primary care population. This is a good example of spectrum mismatch between training and use, and of why local validation is needed even for well-established tools.

Information Beyond the FEV1/FVC Ratio

A spirogram contains far more information than two volumes. Indices from the shape of the curve and from later parts of expiration can signal early airflow impairment in people who do not yet meet the ratio threshold (Hoesterey *et al.*, 2019). Deep learning on spirogram time series can detect COPD and predict its later onset (Mei *et al.*, 2025). In the UK Biobank, a deep convolutional network trained on raw spirograms produced a liability score that identified new genetic loci and improved risk models for COPD, and it also predicted survival and exacerbations (Cosentino *et al.*, 2023). These findings suggest that AI could extract an early-warning signal from tests already being done, which fits the GOLD interest in pre-COPD. The caveat is that such signals have not yet been shown to change management or outcomes.

Chest Computed Tomography

CT-based AI is the most developed AI strategy in COPD (Agusti *et al.*, 2026). It is also the most natural fit for the GOLD 2026 emphasis on lung cancer screening and on CT when findings are discordant.

Detecting COPD on Scans Done for Other Reasons

Tang *et al.* (2020) trained deep residual networks on low-dose CT from 2,589 smokers in a lung cancer screening programme. The networks separated people with spirometrically defined COPD from those without it with an area under the curve (AUC) of 0.87, while mean lung density, a traditional emphysema measure, reached only 0.68. An accompanying commentary saw this as a step towards large-scale case finding (Bibault and Xing, 2020). A weakly supervised deep learning approach detected and staged COPD from CT, and its activation maps pointed to the trachea and lung periphery, regions that standard density and airway measures largely ignore (Sun *et al.*, 2022). In a multicentre retrospective study of 4,106 patients, a machine learning model built on whole-lung inspiratory quantitative CT measurements, called AutoCOPD, was proposed as a practical detection tool (Lin *et al.*, 2025). An editorial on this work warned against seeing such results as magic and stressed that CT findings and spirometry measure different things (Calverley, 2025). Earlier, Gonzalez *et al.* (2018) showed in the COPDGene and ECLIPSE cohorts that deep learning on CT could identify smokers with COPD, stage severity and predict exacerbations and mortality.

Phenotyping Structural Disease

Beyond detection, AI can describe what kind of structural disease is present. Deep learning can automate the visual

Fleischner grading of emphysema, and the automated grade is associated with lung function and mortality (Humphries *et al.*, 2020). Unsupervised learning on CT from the SPIROMICS and MESA Lung studies identified six emphysema subtypes (Angelini *et al.*, 2023). A disease progression model, Subtype and Stage Inference, found two main trajectories in COPDGene: one in which small airway disease and emphysema come first, and one in which large airway wall changes come first. Some smokers with normal spirometry who already sat on these trajectories were more likely to develop COPD later (Young *et al.*, 2020). Machine learning has also been used to characterise COPD subtypes more broadly in COPDGene (Castaldi *et al.*, 2020). AI may help quantify mucus plugs, which are emerging as a possible treatable trait (Agustí *et al.*, 2025). Quantitative CT combined with machine learning can predict pulmonary function (Gawlitza *et al.*, 2019), emergency visits and hospital admissions (Moslemi *et al.*, 2023), and survival (Yun *et al.*, 2021). For diagnosis, the value of phenotyping lies mainly in the grey zone of pre-COPD and PRISm, where structure may tell clinicians more than the ratio does.

Chest Radiographs and Electrocardiograms

Chest radiographs and electrocardiograms (ECGs) are far more common than CT or spirometry, especially in low-resource settings, so even modest accuracy could have wide reach. Schroeder *et al.* (2020) trained a convolutional network on 6,749 two-view radiographs linked to pulmonary function results at a single institution. It predicted obstruction with an AUC of 0.81, better than natural language processing of the radiologists' reports (AUC 0.70). This shows that radiologists' reports often miss COPD that the image contains. On a much larger scale, Ueda *et al.* (2024) used 141,734 radiograph–spirometry pairs from 81,902 patients at five Japanese institutions to estimate FEV1, with a correlation of 0.92 against measured values.

For ECGs, Vaid *et al.* (2026) analysed 208,231 tracings from 18,225 people with COPD and 49,356 matched controls. A convolutional network reached an AUC of 0.80 in internal testing, 0.82 in an external hospital and 0.75 in UK Biobank, and explainability analyses pointed to changes in the P wave. The ECG signal is plausible because COPD affects the right heart and pulmonary circulation. However, the primary outcome was a new COPD diagnosis based on diagnostic codes, not spirometry. This matters because coded diagnoses include both missed and false cases, as noted in the Introduction.

Physiological and Acoustic Signals

Several cheap and portable signals have been studied as possible alternatives or additions to spirometry. Voice recordings analysed with machine learning can separate people with and without COPD in a dedicated dataset (Idrisoglu *et al.*, 2024), and speech features track clinical status during and after exacerbations (Mayr *et al.*, 2025). Convolutional networks applied to electronic nose signals can distinguish COPD from smokers and healthy people (Avian *et al.*, 2022). Forced oscillometry detects small airway dysfunction with little patient effort and is increasingly discussed for early disease detection (Tirelli *et al.*, 2025). The single-breath nitrogen test predicts later lung function decline and COPD (Pistelli *et al.*, 2024). Electrical impedance tomography has revealed abnormal regional ventilation in people whose spirometry is normal (Qu *et al.*, 2025). Most of these studies are small, come from single centres and use case–control designs. They are interesting as research tools for pre-COPD, but none is ready to replace spirometry. GOLD 2026 lists them as being under investigation, which is a fair summary.

Electronic Health Records and Risk Prediction

Health records hold many clues that a person may have COPD: smoking history, repeated chest infections, inhaler prescriptions, occupational exposures, symptom descriptions and imaging reports. AI can search these data and prompt testing (Agusti *et al.*, 2026). Automated tools can extract lung function values from free text in records and improve the identification of COPD for research and care (Akgün *et al.*, 2020). An individualised model built on population data predicts long-term lung function trajectory and the risk of developing COPD (Chen *et al.*, 2020), and machine learning on primary care data from more than 13,000 people predicted a fast decliner phenotype (Nikolaou *et al.*, 2021). Perspectives from China have argued that AI could support early screening in primary health care institutions where specialist resources are thin (Yang, 2024).

The wider prediction literature gives a reason for caution. A systematic review and meta-analysis of machine learning and deep learning models for long-term prognosis in COPD found limited evidence that they outperform existing severity scores, together with a high risk of bias in many studies (Smith *et al.*, 2023). Some models report high discrimination for specific tasks, such as an AUC of 0.87 for predicting severe exacerbations (Zeng *et al.*, 2022) or high accuracy for wearable-based exacerbation prediction in a small cohort (Wu *et al.*, 2021). Systematic reviews of remote monitoring and machine learning show, however, that most work in this area is retrospective and short, with very few studies of clinical impact (Glyde *et al.*, 2024; Chen *et al.*, 2025; Wang *et al.*, 2025). An official American Thoracic Society research statement on home-based monitoring reached a similar conclusion (Khor *et al.*, 2025).

Large Language Models and Foundation Models

Large language models (LLMs) can encode a great deal of clinical knowledge (Singhal *et al.*, 2023), and their possible uses and risks in medicine are now widely discussed (Thirunavukarasu *et al.*, 2023). In COPD diagnosis, plausible uses include reading free-text notes and radiology reports to flag possible cases, summarising a patient's records before a visit, explaining spirometry results to patients and drafting referral letters. In a study of public online questions, chatbot answers were rated as more empathetic and often of higher quality than physicians' answers (Ayers *et al.*, 2023), which hints at a role in patient communication. There are serious risks. LLMs can repeat outdated and race-based medical claims, including claims about lung function (Omiye *et al.*, 2023), and they can produce fluent but wrong statements. Pulmonary and critical care commentators have argued that these tools should assist, not replace, clinical decision-

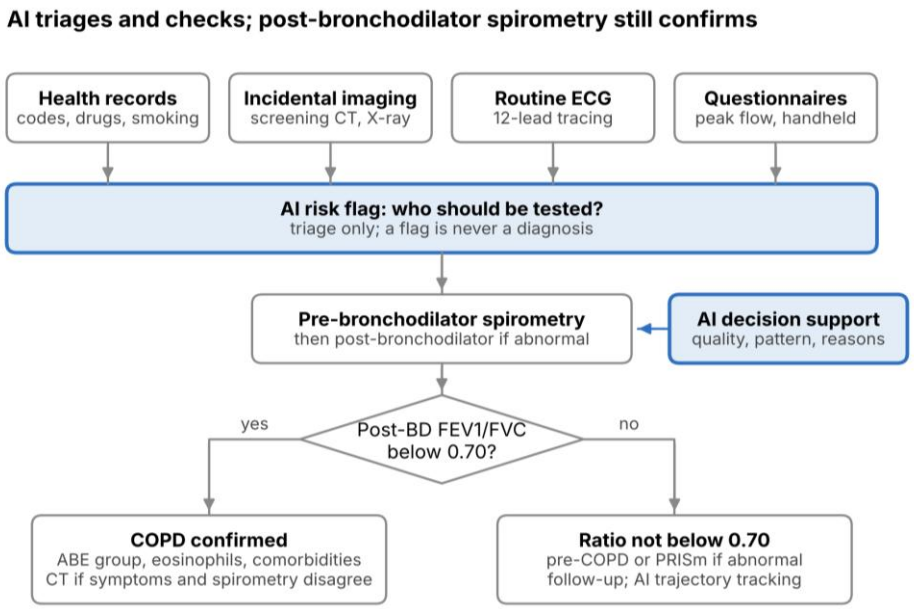


Fig. 1: AI-supported diagnostic pathway under GOLD 2026.
Shaded boxes are AI layers: a risk flag built from data already collected decides who is invited for testing, and decision support helps with test quality and interpretation. Post-bronchodilator spirometry remains the confirmatory step.
BD, bronchodilator; CT, computed tomography; ECG, electrocardiogram; PRISm, preserved ratio impaired spirometry.

making (Friedman *et al.*, 2024). The GOLD perspective also notes that general-purpose foundation models do not always beat specialised models on specific tasks such as COPD diagnosis (Agusti *et al.*, 2026). For now, there is no published evidence that an LLM improves COPD diagnosis in practice.

Bringing the Pieces Together

The studies above use very different data, but together they suggest a coherent role for AI that respects GOLD 2026. Routine data sources can generate a risk flag that prompts testing. AI can then help ensure that spirometry is of good quality and is read correctly. Finally, imaging and longitudinal data can help classify people who do not meet the ratio but are clearly not healthy. Fig. 1 shows this layered pathway. In each layer, AI supports a decision, while post-bronchodilator spirometry remains the confirmatory test.

EVIDENCE GAPS AND CRITICAL APPRAISAL

The results described above are encouraging, but they need to be read with care. Across medicine, systematic reviews have found that deep learning often matches clinicians in diagnostic accuracy on paper, yet few studies are prospective, few are externally validated, and many are poorly reported (Liu *et al.*, 2019; Nagendran *et al.*, 2020). The COPD literature shares these problems and adds some of its own. The main gaps are summarised in Table 3 and discussed below.

Table 3: Main evidence gaps and the study designs needed to close them.

Gap	Why it matters under GOLD 2026	Typical example in current studies	Design that would close it
Reference standard mismatch	GOLD requires post-bronchodilator obstruction plus clinical context	Diagnostic codes or pre-bronchodilator values as labels	Labels from quality-assured post-bronchodilator spirometry
Spectrum bias	Case finding happens in mixed, symptomatic populations	Clear cases versus healthy controls	Consecutive enrolment in intended setting with full verification
Missing calibration and thresholds	Thresholds decide how many people are sent for testing	AUC reported alone	Calibration plots, decision curves, number needed to test
Limited external validation	Scanners, coding and populations differ between sites	Single-centre development	Multi-site temporal and geographic validation
Equity	Labels may encode access to care and race-based equations	No subgroup results	Performance by sex, age, ethnicity, income and setting
Few data from low- and middle-income countries	Most COPD burden lies there	Cohorts from high-income countries	Validation in South Asia, Africa and Latin America
No outcome evidence	Detection is useful only if it improves care	Accuracy studies only	Pragmatic randomised or stepped-wedge trials
Weak reporting	Hidden bias cannot be assessed	Missing data and sample size not described	Adherence to TRIPOD+AI, DECIDE-AI, SPIRIT-AI and CONSORT-AI

Reference Standards That Do Not Match GOLD

A model can only learn the label it is given. For COPD, the correct label under GOLD 2026 is persistent post-bronchodilator obstruction in a person with compatible symptoms and exposures. Many studies use something else. Some use diagnostic codes, which carry the underdiagnosis and misdiagnosis rates described earlier (Diab *et al.*, 2018; Vaid *et al.*, 2026). Others use pre-bronchodilator spirometry, a fixed ratio without regard to symptoms, or composite definitions that also require a low FEV1 (Tang *et al.*, 2020). Some use expert panels with access to full records, which is a strong standard but not the same as the GOLD criteria (Sunjaya *et al.*, 2025). Each choice changes the meaning of a reported AUC. A model trained on coded diagnoses partly learns who gets diagnosed, which may track health care contact, age and sex as much as lung disease. Studies should state exactly which definition they used and, where possible, report results against post-bronchodilator spirometry.

Spectrum Bias and Case–Control Designs

Many development studies compare clear cases with clear controls. This inflates apparent accuracy, because the hard middle ground of mild disease, PRISm, asthma and heart failure is left out. The primary care validation of spirometry software illustrates the issue: a tool trained in secondary care, where normal tests are relatively rare, had to work in a population with many normal and borderline results (Sunjaya *et al.*, 2025). A tool intended for case finding must be tested in the population where case finding happens, which is mostly symptomatic or exposed adults in primary care. The CAPTURE study, although not about AI, shows what such an evaluation looks like: thousands of consecutive patients, all tested with spirometry, so that both sensitivity and specificity are measured without verification bias (Martinez *et al.*, 2023).

Calibration, Thresholds and Clinical Utility

Discrimination, usually reported as an AUC, says whether a model ranks people correctly. It does not say whether the predicted risks are accurate, nor where a threshold should be set. For triage, calibration and the choice of threshold determine how many people are sent for spirometry and how many cases are missed. These figures are rarely reported in COPD AI studies. A model with an AUC of 0.80 could send a manageable number of people for testing in one clinic and overwhelm another, depending on prevalence. Decision curve analysis and estimates of the number needed to test per new diagnosis would make results much easier to use. For comparison, the CAPTURE approach needed about 81 screenings to find one person with treatable undiagnosed COPD, a figure that health planners can work with directly.

Generalisability and Dataset Shift

Deep learning models can pick up shortcuts. A classic example showed that a pneumonia model performed worse in new hospitals partly because it had learned features linked to the hospital rather than to the disease (Zech *et al.*, 2018). Performance also drifts when scanners, protocols, coding practices or patient populations change, a problem known as dataset shift (Finlayson *et al.*, 2021). The external validation of a widely deployed proprietary sepsis model found much poorer performance than its developers had reported (Wong *et al.*, 2021). In COPD imaging, CT reconstruction kernels, slice thickness and radiation dose all affect density measures, and low-dose screening CT differs from diagnostic CT. Most COPD models have been built in North American, European or East Asian cohorts. Very few have been tested in South Asia or sub-Saharan Africa, where biomass exposure, tuberculosis sequelae and early-life factors shape a different form of disease (Stolz *et al.*, 2022; Adeloye *et al.*, 2022).

Equity and the Problem of Reference Equations

Algorithms can reproduce and amplify inequity. A widely used commercial risk tool in the United States assigned lower risk to Black patients than to equally sick White patients, because it used health care costs as a proxy for need (Obermeyer *et al.*, 2019). The same mechanism could operate in COPD if models are trained on coded diagnoses, which depend on access to care. Lung function interpretation adds a specific problem. Many historical datasets were labelled using race-specific reference equations, which can make the same measured value count as normal in one group and abnormal in another (Bhakta *et al.*, 2023). An AI model trained on such labels will inherit that pattern, even if race is never an explicit input. Moves towards race-neutral equations (Bowerman *et al.*, 2023) mean that models may need to be re-labelled and re-validated. LLMs carry a related risk because their training text includes outdated race-based claims about lung function (Omiye *et al.*, 2023). Sex differences also deserve attention, since women are more likely to be underdiagnosed (Diab *et al.*, 2018), and models should report performance by sex, age group, ethnicity and socioeconomic position.

The Missing Outcome Evidence

No randomised trial has yet shown that an AI tool for COPD diagnosis improves patient outcomes, reduces costs or even increases the number of correct diagnoses in routine care. The prognostic literature has the same gap: models have not been shown to beat existing scores (Smith *et al.*, 2023), and remote monitoring has limited evidence of benefit on hard outcomes (Khor *et al.*, 2025). The GOLD Science Committee lists limited prospective validation and uncertain clinical benefit as its first two concerns (Agusti *et al.*, 2026). The UCAP trial is an important model here, because it showed that case finding followed by treatment changes outcomes (Aaron *et al.*, 2024b). The open question is whether AI can find the same patients more efficiently, or find different patients who would also benefit.

Reporting Quality

Reporting standards now exist for every stage of the AI lifecycle. TRIPOD+AI covers prediction models (Collins *et al.*, 2024), PROBAST supports risk-of-bias assessment (Wolff *et al.*, 2019), DECIDE-AI guides early clinical evaluation (Vasey *et al.*, 2022), SPIRIT-AI covers trial protocols (Cruz Rivera *et al.*, 2020) and CONSORT-AI covers trial reports (Liu *et al.*, 2020). Few COPD AI studies cite these standards. Common omissions include how missing data were handled, how the sample size was chosen,

how thresholds were selected, whether the test set was truly independent, and whether code and models are available. Better reporting would not fix weak designs, but it would make weaknesses visible.

A Conceptual Risk: Labelling Without Benefit

GOLD 2026 and the forthcoming 2027 report widen the circle of people of interest to include pre-COPD, PRISm and people with structural abnormalities on CT (Han *et al.*, 2021; Bhatt *et al.*, 2025; Healio, 2026). AI tools that are very good at finding subtle abnormalities could label large numbers of people as abnormal. For many of these states, there is no proven treatment beyond smoking cessation and general measures. Detection is valuable only if it leads to action that helps, and the risk of overdiagnosis, anxiety and unnecessary inhaler use is real (Calverley, 2025). This is not an argument against AI, but it is an argument for linking every AI detection pathway to a clear clinical response and for measuring what happens to the people it flags (Kelly *et al.*, 2019).

IMPLEMENTATION PATHWAYS

The gap between a promising model and a working service is wide. Frameworks such as NASSS, which looks at non-adoption, abandonment, scale-up, spread and sustainability, show that technology projects in health care fail more often because of the condition, the organisation, the workforce and the wider system than because of the technology itself (Greenhalgh *et al.*, 2017). This section proposes how AI could be introduced into COPD diagnosis in steps, with the level of evidence required rising as the level of risk rises. The proposals are summarised in Table 4.

A Staged Model

The study suggests that AI should enter the GOLD 2026 pathway in three stages, ordered by risk and by the strength of current evidence. In the first stage, AI supports the quality and reading of spirometry that a clinician has already decided to perform. The decision remains human, the test remains the reference, and errors are easy to detect. In the second stage, AI flags people who might benefit from spirometry, using data that already exist, such as records, CT scans from lung cancer screening, chest radiographs or ECGs. Here, the main risks are false alarms that waste testing capacity and missed cases that create false reassurance, so local validation and threshold setting are essential. In the third stage, AI contributes to classification and phenotyping beyond the ratio, for example, in pre-COPD, PRISm or disease activity. This stage carries the greatest risk of overdiagnosis and should, for now, remain within research protocols.

This staging reflects a principle stated by the GOLD Science Committee: AI applications should be independently validated, adapted to local contexts with local data and shown to work in representative settings before routine use (Agusti *et al.*, 2026). It also reflects the GOLD 2026 report's own position that AI is an adjunct to clinical judgement and not a substitute for it (Global Initiative for Chronic Obstructive Lung Disease, 2025a).

Primary Care Spirometry Support

The most immediate opportunity is to embed quality control and interpretation support in the spirometers and software already used in primary care. Real-time feedback on acceptability, based on the international standard (Graham *et al.*, 2019), could reduce the number of unusable tests and the number of patients called back. Interpretation support could be offered with an explanation of the reasoning, which has been shown to help clinicians rather than simply override them (Das *et al.*, 2023). Before adoption, each service should check performance on a local sample with expert review, as was done in London primary care (Sunjaya *et al.*, 2025). Training remains essential. Software can flag a poor manoeuvre, but someone still needs to coach the patient to do it better.

Opportunistic Imaging Programmes

Lung cancer screening programmes offer an obvious setting for AI-assisted case finding. Participants are at high risk, CT data are already collected and standardised, and the programmes have an existing structure for follow-up. A workable pathway would apply automated emphysema and airway analysis, or a deep learning COPD score, to every screening scan; report the result in a structured way; and send flagged participants without a diagnosis for spirometry and assessment (Labaki *et al.*, 2024). A prospective pilot without AI has already shown that adding COPD case finding to screening is feasible (Spetrini *et al.*, 2025), and earlier work showed that simple quantitative CT measures can identify many participants with obstruction (Mets *et al.*, 2011). The next step is a prospective study comparing AI-enhanced and

Table 4: A staged pathway for implementing AI in COPD diagnosis under GOLD 2026.

Stage	Use case	Minimum evidence before routine use	Key safeguards	Main outcome to monitor
1. Test support	Spirometry quality control and interpretation	External validation in the same care level; local check against expert review	Clinician sign-off; explanation of output; operator training	Share of acceptable tests; agreement with expert reading
2a. Imaging triage	COPD flag on lung cancer screening CT	Multi-site validation on screening protocols; prospective pilot	Structured reporting; spirometry confirmation for every flag	New diagnoses per 100 scans; time to diagnosis
2b. Record and routine-test triage	Risk lists from records, radiographs or ECGs	Local calibration and threshold setting; subgroup performance	Transparent reasons for flags; equity audit; capacity planning	Number needed to test per diagnosis; missed cases
3. Beyond the ratio	Pre-COPD, PRISm and activity phenotyping	Prospective cohort evidence that flagged groups benefit from action	Research protocols only; avoid labelling without a care plan	Outcomes, treatment changes and harms in flagged people

standard reporting in terms of new diagnoses, treatment started and patient experience. Similar logic could later apply to diagnostic chest CT and chest radiographs taken for other reasons.

Record-Based Triage

Health record triage is cheap to run once built, but hard to get right. Data quality varies widely between practices and countries, and coding habits change over time. A sensible approach starts with simple, transparent rules, such as smoking history plus repeated respiratory prescriptions without spirometry, and adds machine learning only where it clearly improves yield on local data. Automated extraction of lung function values from free text can help identify people who were tested but never labelled, or labelled but never tested (Akgün *et al.*, 2020). Results should be fed to clinicians as a short list of people to invite for spirometry, with the reason for each flag. Model cards or fact labels that state what the model does, on whom it was tested and where it should not be used help clinicians use such lists sensibly (Sendak *et al.*, 2020).

Governance, Regulation and Monitoring

AI tools that influence diagnosis are medical devices in most jurisdictions. In the European Union, the Artificial Intelligence Act classifies such systems as high risk and requires risk management, data governance, human oversight and post-market monitoring (European Parliament and Council of the European Union, 2024). Additionally, with how AI medical devices are regulated in the EU, each model upgrade for an AI system may require a fresh clearance for the device, making the AI adoption more complicated and time-consuming. The World Health Organization has set out ethical principles for AI in health, including protection of autonomy, transparency, accountability, inclusiveness and sustainability (World Health Organization, 2021). In practice, health services need a named clinical owner for each tool, a record of the version in use, routine audit of flagged and missed cases, and a plan for responding to dataset shift, for example after a change of CT scanner or record system (Finlayson *et al.*, 2021). Audits should include performance by sex, age and socioeconomic group. Patients should be told when AI has contributed to a decision about their care.

Low- and Middle-Income Settings

Most people with COPD live in low- and middle-income countries, where spirometry is often unavailable and where biomass smoke, outdoor pollution, tuberculosis and poor early-life lung growth play a large part (Buist *et al.*, 2007; Adeloye *et al.*, 2022; Stolz *et al.*, 2022). In India and similar settings, AI is sometimes presented as a way to bypass spirometry altogether. The evidence does not support that yet. A more realistic near-term role is to make scarce spirometry go further, by prioritising who is tested and by supporting less experienced staff to perform and read tests well. Tools based on chest radiographs or ECGs are attractive because these tests are widely available, but they must be validated locally against spirometry before use. The GOLD perspective rightly asks that AI tools be deployable in resource-limited settings with minimal technical requirements (Agusti *et al.*, 2026), and perspectives from primary care in China make a similar case (Yang, 2024). Local data collection, open models that can be retrained and partnerships with national respiratory programmes will matter more than imported commercial products.

FUTURE RESEARCH AGENDA

The evidence reviewed here points to a short list of research priorities. They are ordered roughly from what could be done now to what will take longer.

The first priority is shared, well-labelled data. The field needs multi-site datasets in which every participant has quality-assured pre- and post-bronchodilator spirometry, symptom scores, exposure history and, where available, CT, radiographs and ECGs. These datasets should include primary care populations and should deliberately include people from South Asia, sub-Saharan Africa and Latin America, as well as people exposed to biomass smoke rather than tobacco. Labels should follow GOLD definitions and race-neutral interpretation, and they should be documented so that models can be re-labelled when definitions change.

The second priority is honest external validation. Every model that claims readiness for clinical use should be tested on data from a different time period, a different site and, ideally, a different country, with calibration, decision curves and subgroup results reported. Head-to-head comparisons with simple tools such as CAPTURE, peak flow or rule-based record searches are needed, because AI is only worth its cost if it adds value over these.

The third priority is prospective and pragmatic trials. The UCAP trial showed that case finding followed by treatment can improve outcomes (Aaron *et al.*, 2024b). A natural next step is a trial in which practices or screening programmes are randomised, or introduced in stepped-wedge fashion, to AI-supported versus usual case finding, with outcomes such as new spirometry-confirmed diagnoses, time to diagnosis, appropriate treatment started, inappropriate inhaler use stopped, health care use and patient experience. Such trials should follow SPIRIT-AI and CONSORT-AI (Cruz Rivera *et al.*, 2020; Liu *et al.*, 2020), and earlier feasibility work should follow DECIDE-AI (Vasey *et al.*, 2022). For spirometry support, trials should test whether AI helps primary care clinicians to interpret results more accurately, not only whether the software is accurate on its own.

The fourth priority is longitudinal modelling that matches the GOLD concepts of trajectory and activity. Lung function trajectories differ widely from early life onwards (Melén *et al.*, 2024), and features from raw spirograms appear to carry information about future risk (Cosentino *et al.*, 2023). Models that combine repeated spirometry, imaging and record data could help define which people with pre-COPD or PRISm progress, and therefore who should be offered closer follow-up or trials of treatment.

The fifth priority is research on how clinicians and patients use AI outputs. Explanations, the wording of flags, the way uncertainty is shown and the handling of disagreement between AI and clinician all influence whether a tool helps or harms. Studies should also measure workload, because a tool

that adds clicks or alerts in an already busy clinic may be abandoned, however accurate it is (Agusti *et al.*, 2026).

Finally, the GOLD report itself will keep changing. The 2027 report is due to be published shortly after this review and will place more weight on CT and on several defined groups within pre-COPD (Healio, 2026). AI research should be designed so that it can adapt to such changes, for example by storing raw data rather than only derived categories.

CONCLUSION

GOLD 2026 keeps post-bronchodilator spirometry at the heart of COPD diagnosis while opening the door to targeted case finding, opportunistic imaging, attention to early disease and a first formal discussion of AI (Global Initiative for Chronic Obstructive Lung Disease, 2025a). This creates clear tasks for AI: deciding who should be tested, improving how tests are performed and read, and adding structural and longitudinal information where the ratio is not enough. The evidence is strongest for spirometry interpretation software and for deep learning on CT, and promising but early for chest radiographs, ECGs, voice, other signals and health records. Across all modalities, the same weaknesses recur: labels that do not match GOLD, selected study populations, missing calibration, limited external validation, little data from the countries with the greatest burden and almost no evidence of impact on patients.

The practical conclusion is that AI should, for now, be used as a triage and quality layer that sends the right people to good-quality spirometry, not as a replacement for it. Health systems that adopt AI tools should validate them locally, set thresholds with their own capacity in mind, audit performance across groups and monitor for drift. Researchers should move from accuracy studies to pragmatic trials with patient-relevant outcomes. If these steps are taken, AI could help close the long-standing gap between the number of people who have COPD and the number who know they have it, which is the problem that GOLD 2026 has placed at the centre of its strategy.

ORCID iD

Phalguni Deswal: 0009-0006-1748-9641

Grant Support Details

The present research did not receive any financial support to conduct the research.

Conflict of Interest

The author declares that there is no conflict of interest regarding the publication of this manuscript. In addition, the ethical issues, including plagiarism, informed consent, misconduct, data fabrication and/ or falsification, double publication and/or submission, and redundancy, have been completely observed by the author.

REFERENCES

1) Aaron, S.D., Montes de Oca, M., Celli, B., *et al.* (2024a) 'Early diagnosis and treatment of chronic obstructive pulmonary disease: the costs and benefits of case finding', *American Journal of Respiratory and Critical Care Medicine*, 209, pp. 928-937. https://doi.org/10.1164/rccm.202311-2120PP

2) Aaron, S.D., Vandemheen, K.L., Whitmore, G.A., *et al.* (2024b) 'Early diagnosis and treatment of COPD and asthma – a randomized, controlled trial', *New England Journal of Medicine*, 390, pp. 2061-2073. https://doi.org/10.1056/NEJMoa2401389

3) Adeloye, D., Song, P., Zhu, Y., *et al.* (2022) 'Global, regional, and national prevalence of, and risk factors for, chronic obstructive pulmonary disease (COPD) in 2019: a systematic review and modelling analysis', *The Lancet Respiratory Medicine*, 10(5), pp. 447-458. https://doi.org/10.1016/S2213-2600(21)00511-7

4) Agusti, A. and Faner, R. (2019) 'Lung function trajectories in health and disease', *The Lancet Respiratory Medicine*, 7, pp. 358-364. https://doi.org/10.1016/S2213-2600(18)30529-0

5) Agustí, A., Celli, B.R., Criner, G.J., *et al.* (2023) 'Global Initiative for Chronic Obstructive Lung Disease 2023 report: GOLD executive summary', *European Respiratory Journal*, 61(4), 2300239. https://doi.org/10.1183/13993003.00239-2023

6) Agustí, A., Aliberti, S., Blasi, F., *et al.* (2025) 'Occluding mucous airway plugs in patients with obstructive lung diseases: a new treatable trait?', *ERJ Open Research*, 11, 00793-2024. https://doi.org/10.1183/23120541.00793-2024

7) Agusti, A., Vila, M., Faner, R., *et al.* (2026) 'How artificial intelligence could improve the diagnosis and management of chronic obstructive pulmonary disease: a perspective from GOLD', *American Journal of Respiratory and Critical Care Medicine*, 212(11), pp. 2736-2744. https://doi.org/10.1093/ajrccm/aamag337

8) Akgün, K.M., Sigel, K., Cheung, K.H., *et al.* (2020) 'Extracting lung function measurements to enhance phenotyping of chronic obstructive pulmonary disease (COPD) in an electronic health record using automated tools', *PLoS One*, 15, e0227730. https://doi.org/10.1371/journal.pone.0227730

9) Angelini, E.D., Yang, J., Balte, P.P., *et al.* (2023) 'Pulmonary emphysema subtypes defined by unsupervised machine learning on CT scans', *Thorax*, 78, pp. 1067-1079. https://doi.org/10.1136/thorax-2022-219158

10) Avian, C., Mahali, M.I., Putro, N.A.S., *et al.* (2022) 'Fx-Net and PureNet: convolutional neural network architecture for discrimination of chronic obstructive pulmonary disease from smokers and healthy subjects through electronic nose signals', *Computers in Biology and Medicine*, 148, 105913. https://doi.org/10.1016/j.compbiomed.2022.105913

11) Ayers, J.W., Poliak, A., Dredze, M., *et al.* (2023) 'Comparing physician and artificial intelligence chatbot responses to patient questions posted to a public social media forum', *JAMA Internal Medicine*, 183, pp. 589-596. https://doi.org/10.1001/jamainternmed.2023.1838

12) Baethge, C., Goldbeck-Wood, S. and Mertens, S. (2019) 'SANRA—a scale for the quality assessment of narrative review articles', *Research Integrity and Peer Review*, 4, 5. https://doi.org/10.1186/s41073-019-0064-8

13) Bhakta, N.R., Bime, C., Kaminsky, D.A., *et al.* (2023) 'Race and ethnicity in pulmonary function test interpretation: an official American Thoracic Society statement', *American Journal of Respiratory and Critical Care Medicine*, 207(8), pp. 978-995. https://doi.org/10.1164/rccm.202302-0310ST

14) Bhatt, S.P., Abadi, E., Anzueto, A., *et al.* (2025) 'A multidimensional diagnostic approach for chronic obstructive pulmonary disease', *JAMA*, 333(24), pp. 2164-2175. https://doi.org/10.1001/jama.2025.7358

15) Bibault, J.E. and Xing, L. (2020) 'Screening for chronic obstructive pulmonary disease with artificial intelligence', *The Lancet Digital Health*, 2, pp. e216-e217. https://doi.org/10.1016/S2589-7500(20)30076-5

16) Bowerman, C., Bhakta, N.R., Brazzale, D., *et al.* (2023) 'A race-neutral approach to the interpretation of lung function measurements', *American Journal of Respiratory and Critical Care Medicine*, 207(6), pp. 768-774. https://doi.org/10.1164/rccm.202205-0963OC

17) Buist, A.S., McBurnie, M.A., Vollmer, W.M., *et al.* (2007) 'International variation in the prevalence of COPD (the BOLD Study): a population-based prevalence study', *The Lancet*, 370(9589), pp. 741-750. https://doi.org/10.1016/S0140-6736(07)61377-4

18) Calverley, P.M.A. (2025) 'Do you believe in magic? Computed tomography, artificial intelligence, and chronic obstructive pulmonary disease', *American Journal of Respiratory and Critical Care Medicine*, 211, pp. 1112-1113. https://doi.org/10.1164/rccm.202502-0521ED

19) Castaldi, P.J., Boueiz, A., Yun, J., *et al.* (2020) 'Machine learning characterization of COPD subtypes: insights from the COPDGene study', *Chest*, 157, pp. 1147-1157. https://doi.org/10.1016/j.chest.2019.11.039

20) Cazzola, M., Bajpai, J., Calzetta, L., *et al.* (2026) 'GOLD 2026: transforming COPD management with early intervention, multi-dimensional assessment, and personalized care', *Drugs*, 86, pp. 581-597. https://doi.org/10.1007/s40265-026-02303-3

21) Chen, W., Sin, D.D., FitzGerald, J.M., *et al.* (2020) 'An individualized prediction model for long-term lung function trajectory and risk of COPD in the general population', *Chest*, 157, pp. 547-557. https://doi.org/10.1016/j.chest.2019.09.003

22) Chen, Z., Hao, J., Sun, H., *et al.* (2025) 'Applications of digital health technologies and artificial intelligence algorithms in COPD: systematic review', *BMC Medical Informatics and Decision Making*, 25, 77. https://doi.org/10.1186/s12911-025-02870-7

23) Christenson, S.A., Smith, B.M., Bafadhel, M., *et al.* (2022) 'Chronic obstructive pulmonary disease', *The Lancet*, 399(10342), pp. 2227-2242. https://doi.org/10.1016/S0140-6736(22)00470-6

24) Collins, G.S., Moons, K.G.M., Dhiman, P., *et al.* (2024) 'TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods', *BMJ*, 385, e078378. https://doi.org/10.1136/bmj-2023-078378

25) Cosentino, J., Behsaz, B., Alipanahi, B., *et al.* (2023) 'Inference of chronic obstructive pulmonary disease with deep learning on raw spirograms identifies new genetic loci and improves risk models', *Nature Genetics*, 55, pp. 787-795. https://doi.org/10.1038/s41588-023-01372-4

26) Cruz Rivera, S., Liu, X., Chan, A.W., *et al.* (2020) 'Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI extension', *Nature Medicine*, 26, pp. 1351-1363. https://doi.org/10.1038/s41591-020-1037-7

27) Das, N., Topalovic, M. and Janssens, W. (2018) 'Artificial intelligence in diagnosis of obstructive lung disease: current status and future potential', *Current Opinion in Pulmonary Medicine*, 24, pp. 117-123. https://doi.org/10.1097/MCP.0000000000000459

28) Das, N., Verstraete, K., Stanojevic, S., *et al.* (2020) 'Deep-learning algorithm helps to standardise ATS/ERS spirometric acceptability and usability criteria', *European Respiratory Journal*, 56, 2000603. https://doi.org/10.1183/13993003.00603-2020

29) 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, 2201720. https://doi.org/10.1183/13993003.01720-2022

30) Diab, N., Gershon, A.S., Sin, D.D., *et al.* (2018) 'Underdiagnosis and overdiagnosis of chronic obstructive pulmonary disease', *American Journal of Respiratory and Critical Care Medicine*, 198, pp. 1130-1139. https://doi.org/10.1164/rccm.201804-0621CI

31) European Parliament and Council of the European Union (2024) *Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)*. Official Journal of the European Union, L series, Publications Office of the European Union, Luxembourg. https://eur-lex.europa.eu/eli/reg/2024/1689/oj

32) Finlayson, S.G., Subbaswamy, A., Singh, K., *et al.* (2021) 'The clinician and dataset shift in artificial intelligence', *New England Journal of Medicine*, 385, pp. 283-286. https://doi.org/10.1056/NEJMc2104626

33) Friedman, S.H., Long, K.J., Sexauer, S., *et al.* (2024) 'Role of artificial intelligence in assisting pulmonary and critical care clinical decision-making', *American Journal of Respiratory and Critical Care Medicine*, 210, pp. 662-664. https://doi.org/10.1164/rccm.202402-0331RR

34) Gawlitza, J., Sturm, T., Spohrer, K., *et al.* (2019) 'Predicting pulmonary function testing from quantified computed tomography using machine learning algorithms in patients with COPD', *Diagnostics*, 9, 33. https://doi.org/10.3390/diagnostics9010033

35) Gerstein, E., Bierbrier, J., Whitmore, G.A., *et al.* (2023) 'Impact of undiagnosed chronic obstructive pulmonary disease and asthma on symptoms, quality of life, healthcare use, and work productivity', *American Journal of Respiratory and Critical Care Medicine*, 208, pp. 1271-1282. https://doi.org/10.1164/rccm.202307-1264OC

36) Global Initiative for Chronic Obstructive Lung Disease (2025a) *Global strategy for prevention, diagnosis and management of COPD: 2026 report*. Version 1.3, Global Initiative for Chronic Obstructive Lung Disease, Deer Park, IL, USA. https://goldcopd.org/wp-content/uploads/2026/01/GOLD-REPORT-2026-v1.3-8Dec2025_WMV2.pdf

37) Global Initiative for Chronic Obstructive Lung Disease (2025b) *GOLD report 2026: key changes summary*. Global Initiative for Chronic Obstructive Lung Disease, Deer Park, IL, USA. https://goldcopd.org/wp-content/uploads/2025/11/KEY-CHANGES-GOLD-2026-10Nov2025.pdf

38) Glyde, H.M.G., Morgan, C., Wilkinson, T.M.A., *et al.* (2024) 'Remote patient monitoring and machine learning in acute exacerbations of chronic obstructive pulmonary disease: dual systematic literature review and narrative synthesis', *Journal of Medical Internet Research*, 26, e52143. https://doi.org/10.2196/52143

39) Gonzalez, G., Ash, S.Y., Vegas-Sánchez-Ferrero, G., *et al.* (2018) 'Disease staging and prognosis in smokers using deep learning in chest computed tomography', *American Journal of Respiratory and Critical Care Medicine*, 197, pp. 193-203. https://doi.org/10.1164/rccm.201705-0860OC

40) Graham, B.L., Steenbruggen, I., Miller, M.R., *et al.* (2019) 'Standardization of spirometry 2019 update: an official American Thoracic Society and European Respiratory Society technical statement', *American Journal of Respiratory and Critical Care Medicine*, 200(8), pp. e70-e88. https://doi.org/10.1164/rccm.201908-1590ST

41) Greenhalgh, T., Wherton, J., Papoutsi, C., *et al.* (2017) 'Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies', *Journal of Medical Internet Research*, 19(11), e367. https://doi.org/10.2196/jmir.8775

42) Han, M.K., Agusti, A., Celli, B.R., *et al.* (2021) 'From GOLD 0 to pre-COPD', *American Journal of Respiratory and Critical Care Medicine*, 203(4), pp. 414-423. https://doi.org/10.1164/rccm.202008-3328PP

43) Healio (2026) *2027 GOLD report discusses CT use, nonpharmacological treatments, CVD in COPD*. Healio Pulmonology news report, 9 November 2026, Healio, Thorofare, NJ, USA. https://www.healio.com/news/pulmonology/20261109/2027-gold-report-discusses-ct-use-nonpharmacological-treatments-cvd-in-copd

44) Hoesterey, D., Das, N., Janssens, W., *et al.* (2019) 'Spirometric indices of early airflow impairment in individuals at risk of developing COPD: spirometry beyond FEV1/FVC', *Respiratory Medicine*, 156, pp. 58-68. https://doi.org/10.1016/j.rmed.2019.08.004

45) Humphries, S.M., Notary, A.M., Centeno, J.P., *et al.* (2020) 'Deep learning enables automatic classification of emphysema pattern at CT', *Radiology*, 294, pp. 434-444. https://doi.org/10.1148/radiol.2019191022

46) Idrisoglu, A., Dallora, A.L., Cheddad, A., *et al.* (2024) 'COPDVD: automated classification of chronic obstructive pulmonary disease on a new collected and evaluated voice dataset', *Artificial Intelligence in Medicine*, 156, 102953. https://doi.org/10.1016/j.artmed.2024.102953

47) Kelly, C.J., Karthikesalingam, A., Suleyman, M., *et al.* (2019) 'Key challenges for delivering clinical impact with artificial intelligence', *BMC Medicine*, 17, 195. https://doi.org/10.1186/s12916-019-1426-2

48) Khor, Y.H., Poberezhets, V., Buhr, R.G., *et al.* (2025) 'Assessment of home-based monitoring in adults with chronic lung disease: an official American Thoracic Society research statement', *American Journal of Respiratory and Critical Care Medicine*, 211, pp. 174-193. https://doi.org/10.1164/rccm.202410-2080ST

49) Koo, H.K., Jin, K.N., Kim, D.K., *et al.* (2016) 'Association of incidental emphysema with annual lung function decline and future development of airflow limitation', *International Journal of Chronic Obstructive Pulmonary Disease*, 11, pp. 161-166. https://doi.org/10.2147/COPD.S96809

50) Labaki, W.W., Agusti, A., Bhatt, S.P., *et al.* (2024) 'Leveraging computed tomography imaging to detect chronic obstructive pulmonary disease and concomitant chronic diseases', *American Journal of Respiratory and Critical Care Medicine*, 210, pp. 281-287. https://doi.org/10.1164/rccm.202402-0407PP

51) Li, X., Lin, Y., He, W., *et al.* (2025) 'Enhancing the interpretation of spirometry: joint utilization of n-order adaptive Fourier decomposition and deep learning techniques', *IEEE Transactions on Instrumentation and Measurement*, 74, pp. 1-14. https://doi.org/10.1109/TIM.2025.3557112

52) Lin, F., Zhang, Z., Wang, J., *et al.* (2025) 'AutoCOPD – a novel and practical machine learning model for COPD detection using whole-lung inspiratory quantitative CT measurements: a retrospective,

multicenter study', *eClinicalMedicine*, 82, 103166. https://doi.org/10.1016/j.eclinm.2025.103166
53) Liu, X., Faes, L., Kale, A.U., *et al.* (2019) 'A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis', *The Lancet Digital Health*, 1(6), pp. e271-e297. https://doi.org/10.1016/S2589-7500(19)30123-2
54) Liu, X., Cruz Rivera, S., Moher, D., *et al.* (2020) 'Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension', *Nature Medicine*, 26, pp. 1364-1374. https://doi.org/10.1038/s41591-020-1034-x
55) Martinez, F.J., Agusti, A., Celli, B.R., *et al.* (2022) 'Treatment trials in young patients with chronic obstructive pulmonary disease and pre-chronic obstructive pulmonary disease patients: time to move forward', *American Journal of Respiratory and Critical Care Medicine*, 205, pp. 275-287. https://doi.org/10.1164/rccm.202107-1663SO
56) Martinez, F.J., Han, M.K., Lopez, C., *et al.* (2023) 'Discriminative accuracy of the CAPTURE tool for identifying chronic obstructive pulmonary disease in US primary care settings', *JAMA*, 329(6), pp. 490-501. https://doi.org/10.1001/jama.2023.0128
57) Mayr, W., Triantafyllopoulos, A., Batliner, A., *et al.* (2025) 'Assessing the clinical and functional status of COPD patients using speech analysis during and after exacerbation', *International Journal of Chronic Obstructive Pulmonary Disease*, 20, pp. 137-147. https://doi.org/10.2147/COPD.S480842
58) Mei, S., Li, X., Zhou, Y., *et al.* (2025) 'Deep learning for detecting and early predicting chronic obstructive pulmonary disease from spirogram time series', *npj Systems Biology and Applications*, 11, 18. https://doi.org/10.1038/s41540-025-00489-y
59) Melén, E., Faner, R., Allinson, J.P., *et al.* (2024) 'Lung-function trajectories: relevance and implementation in clinical practice', *The Lancet*, 403, pp. 1494-1503. https://doi.org/10.1016/S0140-6736(24)00016-3
60) Mets, O.M., Buckens, C.F.M., Zanen, P., *et al.* (2011) 'Identification of chronic obstructive pulmonary disease in lung cancer screening computed tomographic scans', *JAMA*, 306(16), pp. 1775-1781. https://doi.org/10.1001/jama.2011.1531
61) Moslemi, A., Makimoto, K., Tan, W.C., *et al.* (2023) 'Quantitative CT lung imaging and machine learning improves prediction of emergency room visits and hospitalizations in COPD', *Academic Radiology*, 30, pp. 707-716. https://doi.org/10.1016/j.acra.2022.05.009
62) Nagendran, M., Chen, Y., Lovejoy, C.A., *et al.* (2020) 'Artificial intelligence versus clinicians: systematic review of design, reporting standards, and claims of deep learning studies', *BMJ*, 368, m689. https://doi.org/10.1136/bmj.m689
63) Nikolaou, V., Massaro, S., Garn, W., *et al.* (2021) 'Fast decliner phenotype of chronic obstructive pulmonary disease (COPD): applying machine learning for predicting lung function loss', *BMJ Open Respiratory Research*, 8, e000980. https://doi.org/10.1136/bmjresp-2021-000980
64) Obermeyer, Z., Powers, B., Vogeli, C., *et al.* (2019) 'Dissecting racial bias in an algorithm used to manage the health of populations', *Science*, 366(6464), pp. 447-453. https://doi.org/10.1126/science.aax2342
65) Omiye, J.A., Lester, J.C., Spichak, S., *et al.* (2023) 'Large language models propagate race-based medicine', *npj Digital Medicine*, 6, 195. https://doi.org/10.1038/s41746-023-00939-z
66) Pistelli, F., Sherrill, D.L., Di Pede, F., *et al.* (2024) 'Single breath nitrogen test as predictor of lung function decline and COPD over an 8-year follow-up', *Pulmonology*, 30, pp. 546-554. https://doi.org/10.1016/j.pulmoe.2022.09.001
67) Qu, S., Feng, E., Dong, D., *et al.* (2025) 'Early screening of lung function by electrical impedance tomography in people with normal spirometry reveals unrecognized pathological features', *Nature Communications*, 16, 622. https://doi.org/10.1038/s41467-024-55505-2
68) Quanjer, P.H., Stanojevic, S., Cole, T.J., *et al.* (2012) 'Multi-ethnic reference values for spirometry for the 3–95-yr age range: the global lung function 2012 equations', *European Respiratory Journal*, 40(6), pp. 1324-1343. https://doi.org/10.1183/09031936.00080312
69) Regan, E.A., Lynch, D.A., Curran-Everett, D., *et al.* (2015) 'Clinical and radiologic disease in smokers with normal spirometry', *JAMA Internal Medicine*, 175(9), pp. 1539-1549. https://doi.org/10.1001/jamainternmed.2015.2735
70) Ritchie, A.I., Donaldson, G.C., Hoffman, E.A., *et al.* (2024) 'Structural predictors of lung function decline in young smokers with normal spirometry', *American Journal of Respiratory and Critical Care Medicine*, 209(10), pp. 1208-1218. https://doi.org/10.1164/rccm.202307-1203OC
71) Schroeder, J.D., Bigolin Lanfredi, R., Li, T., *et al.* (2020) 'Prediction of obstructive lung disease from chest radiographs via deep learning trained on pulmonary function data', *International Journal of Chronic Obstructive Pulmonary Disease*, 15, pp. 3455-3466. https://doi.org/10.2147/COPD.S279850
72) Sendak, M.P., Gao, M., Brajer, N., *et al.* (2020) 'Presenting machine learning model information to clinical end users with model facts labels', *npj Digital Medicine*, 3, 41. https://doi.org/10.1038/s41746-020-0253-3
73) Singh, D., Stockley, R., Anzueto, A., *et al.* (2025) 'GOLD Science Committee recommendations for the use of pre- and post-bronchodilator spirometry for the diagnosis of COPD', *European Respiratory Journal*, 65(2), 2401603. https://doi.org/10.1183/13993003.01603-2024
74) Singhal, K., Azizi, S., Tu, T., *et al.* (2023) 'Large language models encode clinical knowledge', *Nature*, 620, pp. 172-180. https://doi.org/10.1038/s41586-023-06291-2
75) Smith, L.A., Oakden-Rayner, L., Bird, A., *et al.* (2023) 'Machine learning and deep learning predictive models for long-term prognosis in patients with chronic obstructive pulmonary disease: a systematic review and meta-analysis', *The Lancet Digital Health*, 5, pp. e872-e881. https://doi.org/10.1016/S2589-7500(23)00177-2
76) Spetrini, R., Pikman, P., Kang, V., *et al.* (2025) 'Prospective COPD case finding in a lung cancer screening program: a pilot study', *Chronic Obstructive Pulmonary Diseases: Journal of the COPD Foundation*, 12, pp. 411-418. https://doi.org/10.15326/jcopdf.2025.0636
77) Stanojevic, S., Kaminsky, D.A., Miller, M.R., *et al.* (2022) 'ERS/ATS technical standard on interpretive strategies for routine lung function tests', *European Respiratory Journal*, 60(1), 2101499. https://doi.org/10.1183/13993003.01499-2021
78) Stolz, D., Mkorombindo, T., Schumann, D.M., *et al.* (2022) 'Towards the elimination of chronic obstructive pulmonary disease: a Lancet Commission', *The Lancet*, 400(10356), pp. 921-972. https://doi.org/10.1016/S0140-6736(22)01273-9
79) Sun, J., Liao, X., Yan, Y., *et al.* (2022) 'Detection and staging of chronic obstructive pulmonary disease using a computed tomography-based weakly supervised deep learning approach', *European Radiology*, 32, pp. 5319-5329. https://doi.org/10.1007/s00330-022-08632-7
80) Sunjaya, A., Edwards, G.D., Harvey, J., *et al.* (2025) 'Validation of artificial intelligence spirometry diagnostic support software in primary care: a blinded diagnostic accuracy study', *ERJ Open Research*, 11(5), 00116-2025. https://doi.org/10.1183/23120541.00116-2025
81) Tang, L.Y.W., Coxson, H.O., Lam, S., *et al.* (2020) 'Towards large-scale case-finding: training and validation of residual networks for detection of chronic obstructive pulmonary disease using low-dose CT', *The Lancet Digital Health*, 2, pp. e259-e267. https://doi.org/10.1016/S2589-7500(20)30064-9
82) Thirunavukarasu, A.J., Ting, D.S.J., Elangovan, K., *et al.* (2023) 'Large language models in medicine', *Nature Medicine*, 29, pp. 1930-1940. https://doi.org/10.1038/s41591-023-02448-8
83) Tirelli, C., Mira, S., Italia, M., *et al.* (2025) 'Applications of forced oscillatory technique in obstructive and restrictive pulmonary diseases: a concise state of the art', *Journal of Clinical Medicine*, 14, 5718. https://doi.org/10.3390/jcm14165718
84) 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
85) Topol, E.J. (2019) 'High-performance medicine: the convergence of human and artificial intelligence', *Nature Medicine*, 25, pp. 44-56. https://doi.org/10.1038/s41591-018-0300-7
86) 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, pp. e580-e588. https://doi.org/10.1016/S2589-7500(24)00113-4
87) Vaid, A., Sharma, J., Jiang, J., *et al.* (2026) 'Automated diagnosis of chronic obstructive pulmonary disease using deep learning applied to

electrocardiograms', *eBioMedicine*, 123, 106066. https://doi.org/10.1016/j.ebiom.2025.106066

88) Vasey, B., Nagendran, M., Campbell, B., *et al.* (2022) 'Reporting guideline for the early-stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI', *Nature Medicine*, 28, pp. 924-933. https://doi.org/10.1038/s41591-022-01772-9

89) Venkatesan, P. (2025) 'GOLD COPD report: 2026 update', *The Lancet Respiratory Medicine*, 14, pp. e14-e15. https://doi.org/10.1016/S2213-2600(25)00432-1

90) Vila, M., Sisó-Almirall, A., Ocaña, A., *et al.* (2025) 'Prevalence, diagnostic accuracy, and healthcare utilization patterns in patients with COPD in primary healthcare: a population-based study', *npj Primary Care Respiratory Medicine*, 35, 17. https://doi.org/10.1038/s41533-025-00419-9

91) Wang, M., Li, L., Feng, M., *et al.* (2025) 'Advances in artificial intelligence applications for the management of chronic obstructive pulmonary disease', *Frontiers in Medicine*, 12, 1685254. https://doi.org/10.3389/fmed.2025.1685254

92) Wang, Y., Li, Q., Chen, W., *et al.* (2022a) 'Deep learning-based analytic models based on flow-volume curves for identifying ventilatory patterns', *Frontiers in Physiology*, 13, 824000. https://doi.org/10.3389/fphys.2022.824000

93) Wang, Y., Li, Y., Chen, W., *et al.* (2022b) 'Deep learning for automatic upper airway obstruction detection by analysis of flow-volume curve', *Respiration*, 101, pp. 841-850. https://doi.org/10.1159/000524598

94) Wang, Y., Li, Y., Chen, W., *et al.* (2022c) 'Deep learning for spirometry quality assurance with spirometric indices and curves', *Respiratory Research*, 23, 98. https://doi.org/10.1186/s12931-022-02014-9

95) Wijnant, S.R.A., De Roos, E., Kavousi, M., *et al.* (2020) 'Trajectory and mortality of preserved ratio impaired spirometry: the Rotterdam Study', *European Respiratory Journal*, 55(1), 1901217. https://doi.org/10.1183/13993003.01217-2019

96) Wolff, R.F., Moons, K.G.M., Riley, R.D., *et al.* (2019) 'PROBAST: a tool to assess the risk of bias and applicability of prediction model studies', *Annals of Internal Medicine*, 170(1), pp. 51-58. https://doi.org/10.7326/M18-1376

97) Wong, A., Otles, E., Donnelly, J.P., *et al.* (2021) 'External validation of a widely implemented proprietary sepsis prediction model in hospitalized patients', *JAMA Internal Medicine*, 181(8), pp. 1065-1070. https://doi.org/10.1001/jamainternmed.2021.2626

98) Woodruff, P.G., Barr, R.G., Bleecker, E., *et al.* (2016) 'Clinical significance of symptoms in smokers with preserved pulmonary function', *New England Journal of Medicine*, 374(19), pp. 1811-1821. https://doi.org/10.1056/NEJMoa1505971

99) World Health Organization (2021) *Ethics and governance of artificial intelligence for health: WHO guidance*. World Health Organization, Geneva. https://www.who.int/publications/i/item/9789240029200

100) Wu, C.T., Li, G.H., Huang, C.T., *et al.* (2021) 'Acute exacerbation of a chronic obstructive pulmonary disease prediction system using wearable device data, machine learning, and deep learning: development and cohort study', *JMIR mHealth and uHealth*, 9, e22591. https://doi.org/10.2196/22591

101) Yang, X. (2024) 'Application and prospects of artificial intelligence technology in early screening of chronic obstructive pulmonary disease at primary healthcare institutions in China', *International Journal of Chronic Obstructive Pulmonary Disease*, 19, pp. 1061-1067. https://doi.org/10.2147/COPD.S458935

102) Young, A.L., Bragman, F.J.S., Rangelov, B., *et al.* (2020) 'Disease progression modeling in chronic obstructive pulmonary disease', *American Journal of Respiratory and Critical Care Medicine*, 201, pp. 294-302. https://doi.org/10.1164/rccm.201908-1600OC

103) Yun, J., Cho, Y.H., Lee, S.M., *et al.* (2021) 'Deep radiomics-based survival prediction in patients with chronic obstructive pulmonary disease', *Scientific Reports*, 11, 15144. https://doi.org/10.1038/s41598-021-94535-4

104) Zech, J.R., Badgeley, M.A., Liu, M., *et al.* (2018) 'Variable generalization performance of a deep learning model to detect pneumonia in chest radiographs: a cross-sectional study', *PLoS Medicine*, 15(11), e1002683. https://doi.org/10.1371/journal.pmed.1002683

105) Zeng, S., Arjomandi, M., Tong, Y., *et al.* (2022) 'Developing a machine learning model to predict severe chronic obstructive pulmonary disease exacerbations: retrospective cohort study', *Journal of Medical Internet Research*, 24, e28953. https://doi.org/10.2196/28953