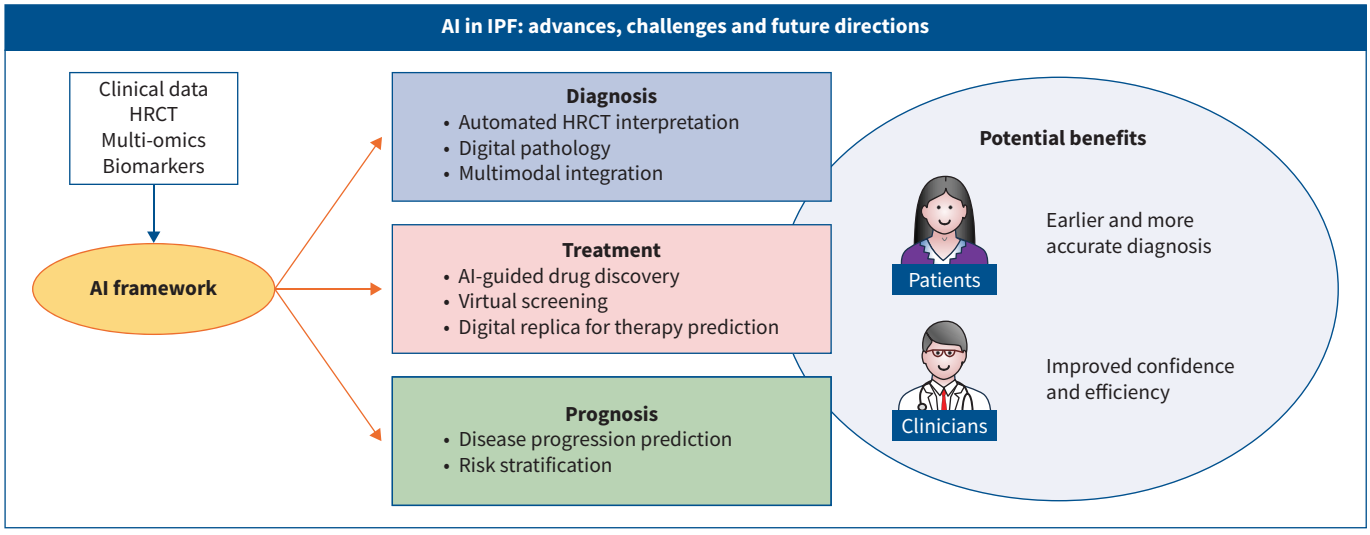
Artificial intelligence in idiopathic pulmonary fibrosis: advances, challenges and future directions

Moisés Selman, Ivette Buendia-Roldan and Annie Pardo



GRAPHICAL ABSTRACT Artificial intelligence (AI) supports timely diagnosis, predicts individual disease trajectory, and drives faster discovery and personalisation of therapies in idiopathic pulmonary fibrosis (IPF), benefiting patients and clinicians. HRCT: high-resolution computed tomography.

https://doi.org/10.1183/13993003.01112-2025 Eur Respir J 2026; 67: 2501112

Artificial intelligence in idiopathic pulmonary fibrosis: advances, challenges and future directions

Moisés Selman [1], **Ivette Buendia-Roldan** [1] and **Annie Pardo** [2]

[1]Instituto Nacional de Enfermedades Respiratorias "Ismael Cosío Villegas", Mexico City, Mexico. [2]Facultad de Ciencias, Universidad Nacional Autónoma de México, Mexico City, Mexico.

Corresponding author: Moisés Selman (mselmanl@yahoo.com.mx)

Shareable abstract (@ERSpublications)
Artificial intelligence may enhance the diagnosis, prognosis and treatment of IPF by integrating clinical, imaging and genomic data, enabling early detection, accurate risk prediction and personalised patient management https://bit.ly/3LM5mQ9

Cite this article as: Selman M, Buendia-Roldan I, Pardo A. Artificial intelligence in idiopathic pulmonary fibrosis: advances, challenges and future directions. *Eur Respir J* 2026; 67: 2501112 [DOI: 10.1183/13993003.01112-2025].

Received: 30 May 2025
Accepted: 16 Nov 2025

Abstract

Idiopathic pulmonary fibrosis (IPF) is a progressive disease of unknown aetiology, characterised by a radiological and/or morphological pattern of usual interstitial pneumonia. Its diagnosis is challenging, and disease progression is often variable and unpredictable. In recent years the introduction of artificial intelligence (AI), particularly machine learning (ML) and deep learning (DL) models, has shown the potential to improve the diagnosis, prognosis and therapeutic strategies for IPF. As part of DL, convolutional neural networks enhance the accuracy of high-resolution computed tomography analysis, facilitating early and precise diagnosis. Likewise, predictive ML and DL models are being developed using clinical, morphological, transcriptional and imaging data to assess disease progression and stratify patients by risk, thereby improving prognosis evaluation. Furthermore, AI-driven drug discovery may optimise treatment strategies by identifying novel therapeutic targets, as recently demonstrated with the discovery of an NCK-interacting kinase inhibitor with strong antifibrotic properties. However, several challenges hamper widespread clinical integration and real-life implementation, including data heterogeneity, model interpretability and the need for robust validation through large-scale, multicentre studies. Future research should prioritise the development of standardised models of AI in large cohorts of IPF patients, combining clinical, imaging, morphological, multi-omics and other data, and enhance model transparency to strengthen clinical confidence. With continued advancements, AI holds potential to improve IPF management, enabling early diagnosis, individualised prognosis and targeted therapy, all aimed at improving patient outcomes. In this review, we explore the evolving role of AI in IPF management, its potential to support clinical decisions and the challenges to its clinical integration.

Introduction

Idiopathic pulmonary fibrosis (IPF) is a progressive and devastating disease, characterised by diagnostic challenges, poor prognosis and a lack of effective treatment [1]. IPF is typically diagnosed by recognising a radiological or histological pattern of usual interstitial pneumonia (UIP). However, confirmation should involve a discussion among a multidisciplinary team, including pulmonologists, radiologists and pathologists, considering clinical data, findings on high-resolution computed tomography (HRCT) and lung biopsy if available.

In recent years, increasing evidence has supported the application of artificial intelligence (AI) in interstitial lung diseases (ILDs), especially in IPF, yielding promising preliminary results.

AI is the capability of computers to execute a variety of complex tasks that mimic human intelligence and cognitive processes, including learning and problem solving. Key elements of AI, machine learning (ML) and deep learning (DL), with their ability to rapidly process and evaluate massive datasets, and identify

complex patterns and correlations, are increasingly being used in various biomedical fields, such as imaging, histopathological analysis, disease diagnosis and drug design, among others [2, 3].

ML is a subdomain of AI that focuses on developing algorithms and models capable of recognising patterns and making decisions without being explicitly programmed for specific tasks [4]. Within ML, DL represents a specialised approach that employs multilayered artificial neural networks to automatically identify complex patterns and relationships from large datasets [5]. A relevant example of DL is convolutional neural networks (CNNs), which have proven particularly effective in image processing, allowing the systematisation of complex analysis tasks with high accuracy (figure 1) [6, 7].

Importantly, defining whether an AI study employs a supervised or unsupervised learning approach is essential for understanding its methodology, interpretability and clinical applicability [7, 8]. Supervised DL is a method in which models are trained using labelled data, *i.e.* datasets where each input (such as an image or clinical profile) has an associated, known output or label. This approach is particularly effective for tasks like classification and disease prediction. For example, in the area of ILD, CNNs can be trained on HRCT scans labelled by expert radiologists to automatically classify patterns consistent with UIP. Likewise, supervised models can use structured clinical variables (such as age and pulmonary function tests (PFTs)) to predict survival or progression, enabling more personalised prognosis.

In contrast, unsupervised DL involves training models on data that lack explicit labels. The goal here is to uncover hidden structures or patterns within the dataset, making it especially important for discovering unknown subtypes or features. Techniques like clustering algorithms can be applied to HRCT scans, gene expression data or other patient-level datasets to group similar patterns without predefined categories. For instance, unsupervised models can identify subgroups of ILD with distinct imaging phenotypes, potentially uncovering new ILD endotypes or treatment-responsive clusters. Thus, clearly specifying the learning type not only ensures transparency and reproducibility but also helps researchers and clinicians assess the scope, limitations and potential impact of the AI-driven findings within the context of biomedical research.

AI-based tools have shown promise in IPF, and table 1 provides a summary of the key learnings from their application in this disease.

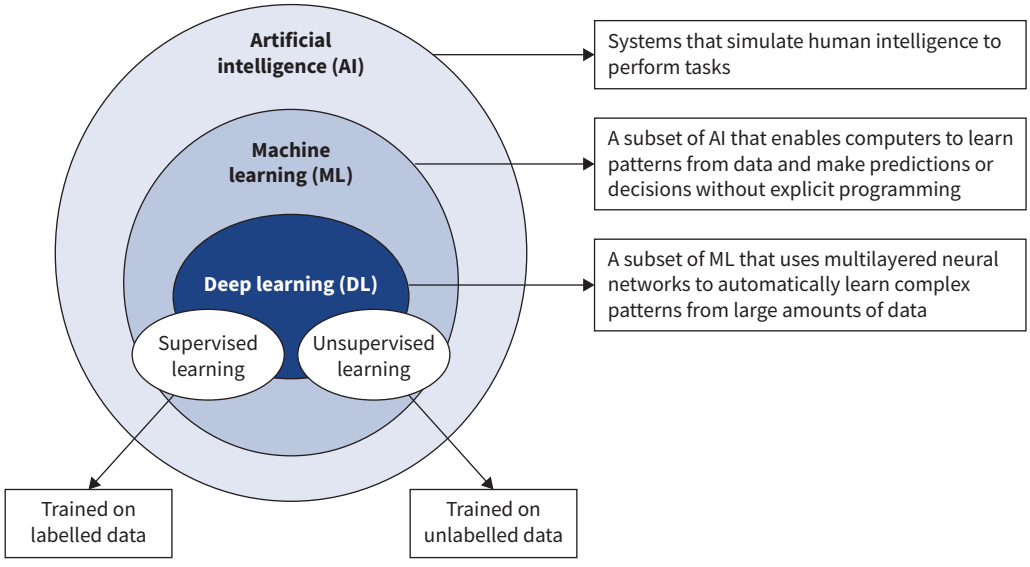


FIGURE 1 Conceptual hierarchy and learning paradigms in artificial intelligence (AI) applied to idiopathic pulmonary fibrosis (IPF). AI encompasses machine learning (ML), which employs algorithms that learn from data to improve performance without explicit programming. Deep learning (DL), a subset of ML, uses multilayered neural networks to capture complex patterns and abstractions. Both ML and DL may be supervised or unsupervised. Supervised learning depends on labelled datasets (*e.g.* high-resolution computed tomography-based usual interstitial pneumonia (UIP) *versus* non-UIP classification for IPF diagnosis), whereas unsupervised learning analyses unlabelled data to reveal hidden structures (*e.g.* discovery of novel IPF imaging subtypes).

TABLE 1 Key learnings from artificial intelligence (AI) in idiopathic pulmonary fibrosis (IPF)

Task	Data modality	Representative methods	Key insights
Diagnosis	HRCT	Convolutional neural networks (*e.g.* Inception-ResNet, SOFIA)	Improved accuracy over visual UIP assessment
	Digital pathology	Convolutional neural networks, computational histopathology models, feature extraction	Detects UIP histological features; supports diagnosis when HRCT is indeterminate
	Genomic tissue classifier	Machine learning model using high-dimensional RNA sequencing data	Differentiates UIP from non-UIP, even in small lung specimens
	Multimodal (radiology, pathology and clinical)	Multimodal fusion networks, late integration deep learning	Combining modalities reduces diagnostic uncertainty
Prognosis	Unimodal approach: HRCT, pathology, clinical, genomic	Deep learning architectures, random forest models	Predicts disease progression, forced vital capacity decline, survival
	Multimodal approach: HRCT, pathology clinical, genomic, proteomic	Data fusion and deep multimodal networks	Enhances prognostic performance; multimodal inputs outperform single modality
Subtype stratification/ molecular classification	Genomics, transcriptomics, proteomics	Unsupervised clustering, machine learning-based classifiers	Identifies molecular subtypes and pathogenic pathways that may inform targeted therapy
Phenotyping	Clinical, HRCT, multi-omics	Unsupervised learning (*e.g.* hierarchical clustering)	Reveals novel IPF phenotypes
Treatment response	Longitudinal clinical data, HRCT	Machine learning on time-series datasets	Predicts responders to antifibrotic therapy; supports earlier treatment decisions
Drug discovery	Genomics, proteomics, knowledge graphs	Network-based AI, graph neural networks, generative models	Identifies new drug targets and repurposes existing compounds efficiently

HRCT: high-resolution computed tomography; UIP: usual interstitial pneumonia.

Nevertheless, this area of research is still in its early stages and several challenges must be overcome before AI-derived technologies can be routinely integrated into clinical practice.

In this review, we examine the emerging role of AI in the diagnosis, prognosis and treatment of IPF, highlighting its potential to assist clinical decision making and improve patient outcomes. We also discuss some challenges that must be addressed to allow its effective integration in the clinical management of this disease. Although we are not specialists in computer science or AI methodology, our intention is to provide clinicians with a clear and practical perspective on its applications and emerging insights in IPF, rather than a technical exposition. Our focus is on translating these advances into clinically meaningful concepts that may inform practice and stimulate further investigation.

Search strategy and selection criteria

We conducted a search on PubMed for articles published in English between January 2010 and March 2025, using the search terms “Idiopathic Pulmonary Fibrosis”, “IPF” combined with “Artificial Intelligence”, “Machine-learning”, “Deep-learning”, “Convolutional Neural Network”. The final list of references was chosen based on their relevance to the focus of this review.

AI in the diagnostic and prognostic evaluation of IPF

Diagnosing and predicting the prognosis of IPF remains a significant clinical challenge due to its complex and heterogeneous behaviour. Factors such as comorbidities, gene variants and environmental exposures further contribute to this variability.

In recent years, AI-driven approaches have been used to explore a range of lung-specific and systemic variables to enhance the diagnosis and prognosis of IPF, including imaging features derived from HRCT, PFTs, clinical parameters and molecular or genetic data.

Fine crackles on auscultation

The presence of bilateral Velcro-like crackles during chest auscultation can indicate the existence of a UIP pattern on HRCT [9]. Notably, these sounds can be detected early in disease progression, and may also be present in asymptomatic individuals and those with normal pulmonary function. These fine crackles can be

detected irrespective of the clinician's experience in performing chest auscultation [9, 10]. In this context, DL models utilising a CNN have been created to facilitate electronic auscultation of lung sounds using digital stethoscopes, enhancing diagnostic precision [11]. A recent study revealed that fine crackles quantitative values (FCQVs), obtained through a supervised ML analysis algorithm, can effectively distinguish lung fields impacted by fibrotic ILD on HRCT scans [12]. Importantly, a high FCQV was significantly associated with the diagnosis of IPF, although the study was performed with a limited number of selected subjects and without independent validation [12].

Pulmonary function tests

A ML algorithm was developed to interpret PFTs, demonstrating improved precision in detecting ILD [13]. In this study, AI improved the detection of lung fibrosis as the primary diagnosis from 42.8% to 72.1%. Instead of depending exclusively on conventional thresholds and indices, this AI model analyses the complete shape of the flow–volume curve and its correlation with other metrics, detecting early indicators of disease that might not be apparent through standard interpretation [13].

In this context, a fully automatic, supervised DL-based multidimensional model able to estimate total lung capacity (TLC) with use of age, sex and chest radiographs was recently developed and validated [14]. The model predicted survival, showing that a greater percentage of TLC estimated by DL was associated with better survival [14].

Histopathology

The pathological diagnosis of IPF/UIP remains challenging due to difficulties in obtaining adequate tissue samples and variability in histopathological interpretation [15]. In 2015, a supervised ML model using high-dimensional RNA sequencing data from lung biopsies was developed, accurately distinguishing IPF/UIP from other ILDs with high sensitivity and specificity [16]. Subsequently, it was demonstrated that this algorithm (Envisia genomic classifier) can differentiate UIP from non-UIP, even in small lung specimens obtained through transbronchial lung biopsies, although with reduced performance [17]. Prospectively, this ML algorithm was able to identified a biopsy-confirmed UIP pattern with 88% specificity and 70% sensitivity in a clinically and geographically diverse group of patients [18, 19]. Also, using an unsupervised ML-based approach, two predominant subpopulations of fibroblasts were identified within fibrotic lung tissues. These fibroblast subtypes exhibited distinct molecular and biological characteristics, suggesting a prognostic utility of their signature genes [20]. High expression of *SPON2* (spondin 2), *FSTL1* (follistatin-like 1), *CCDC80* (coiled-coil domain-containing protein 80), *COL8A1* (collagen type VIII alpha-1 chain) and *FBLN2* (fibulin 2) associated with a poor prognosis.

In another diagnostic approach, self-supervised learning and DL models were also developed to distinguish between UIP and non-UIP patterns based on the frequency of specific morphological features such as dense fibrosis and fibroblastic foci [21]. These models demonstrated high diagnostic performance, achieving an area under the curve (AUC) of 0.86 in the test cohort and 0.90 in the validation cohort.

Likewise, an AI model developed with CNN and supervised learning, trained to detect a range of structural abnormalities in IPF lungs, revealed that a larger area occupied by fibroblastic foci was significantly associated with lower diffusing capacity of the lung for carbon monoxide (D_{LCO}) and predicted worse clinical outcomes. Conversely, a higher number of macrophages correlated with improved survival, suggesting a potential protective role of these cells [22]. Notably, while the number of fibroblastic foci identified by pathologists has been linked to poor prognosis in multiple previous studies [23–25], the application of CNNs enables the detection and quantification of subtle morphological features, offering a more comprehensive and objective evaluation of histological patterns in IPF.

Certainly, there is a fundamental distinction between the use of pathological tissue for genomic classifier development and the application of AI for assisting pathological diagnosis. Although both contribute to improved diagnostic accuracy, their methodologies and purposes differ: the former interrogates gene expression data for molecular characterisation, whereas the latter analyses image-based histological features for pattern recognition (figure 2).

Radiology

AI-driven approaches for diagnosing and predicting IPF have largely focused on HRCT imaging, aiming to identify IPF/UIP patterns that even expert radiologists might overlook.

A common approach is to apply AI at the quantification stage, where algorithms segment or classify CT lung voxels into radiological patterns and compute objective measures of disease extent. These quantitative

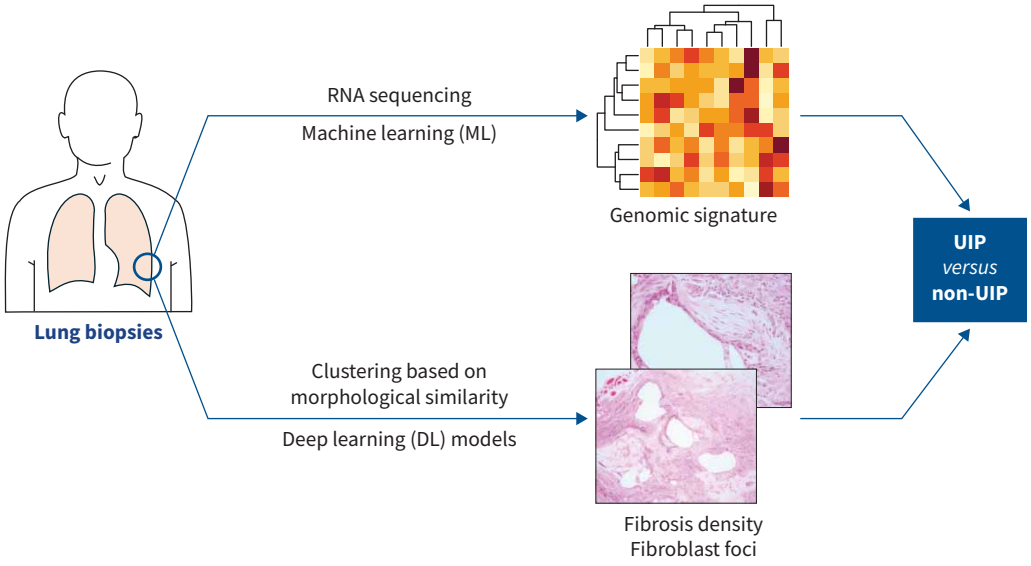


FIGURE 2 Machine learning (ML) and deep learning (DL) approaches for histological diagnosis in idiopathic pulmonary fibrosis (IPF). Biopsy-derived lung tissue can be analysed with a supervised ML model using high-dimensional RNA sequencing data (*e.g.* the Envisia genomic classifier), which accurately differentiates usual interstitial pneumonia (UIP)/IPF from other interstitial lung diseases with high sensitivity and specificity. In parallel, self-supervised DL models distinguish UIP from non-UIP patterns by capturing key morphological features, including dense fibrosis and fibroblastic foci. Together, these molecular and morphological artificial intelligence-based strategies offer complementary diagnostic insights, enhancing accuracy and reproducibility in the evaluation of IPF.

features are then linked to clinical outcomes, serving both diagnostic and prognostic purposes. This framework, using imaging-derived features as biomarkers, forms the basis of widely used tools, and underpins the promise of newer DL approaches, making it central to the development of imaging as a biomarker in ILD.

AI in HRCT-based diagnosis of IPF

In 2018, Walsh *et al.* [26] trained a supervised DL algorithm for rapid and reproducible classification of fibrotic lung disease on HRCT. The specific neural network architecture used in this study was the CNN Inception-ResNet-v2. The study included a training set (n=929), validation set (n=89) and test set (n=139). The algorithm classified cases with greater accuracy and discrimination between UIP *versus* non-UIP and was equivalent to the majority opinion of the expert thoracic radiologist group [26]. In a similar study, a large number of patients (n=1038 for training/validation and n=198 for test performance) with pathologically proven ILD were assigned a label based on histopathological diagnosis (UIP or non-UIP). Results showed that a supervised DL model provides higher diagnostic performance than visual evaluation in predicting UIP/IPF histopathology from CT imaging [27].

Results have been less encouraging when the algorithms have been used in classification of different ILDs. For example, a content-based image retrieval (CBIR) system for chest CT images using two-dimensional-based supervised DL was proposed to improve the diagnostic accuracy of ILD and inter-reader agreement in radiologists with varying experience in these diseases [28]. The overall accuracy of the radiologists improved significantly after applying CBIR in UIP and non-specific interstitial pneumonia, but results with cryptogenic organising pneumonia and chronic hypersensitivity pneumonitis were less satisfactory.

Recently, in 2025, promising findings were reported with the development of a novel DL model, the Parallel Multi-scale Feature Fusion Network (PMFF-Net), which integrates transformer and CNN architectures for ILD classification [29]. The model was evaluated using 3002 HRCT images from 130 patients, achieving diagnostic accuracies of 97.6% for UIP, 82.3% for non-specific interstitial pneumonia, 87.7% for cryptogenic organising pneumonia and 100% for normal scans. Nonetheless, external validation has not yet been performed.

In another approach, a supervised DL classifier was developed using pre-trained three-dimensional ResNet (Residual Network) models, with CT images serving as input and the likelihood of IPF *versus* non-IPF as the output variable [30]. The model was trained on a dataset of over 2000 patients with ILD, approximately 40% with IPF, and underwent external validation in 295 cases. The results demonstrated consistent performance in identifying IPF among ILD cases. This ML software system (named Fibresolve) exhibited a sensitivity of 76.5%, further supporting its potential in the evaluation of ILD patients with a non-definitive UIP pattern and suspected IPF [31]. This algorithm displayed higher sensitivity although similar specificity compared to an expert clinical panel that had access to imaging, clinical, demographic and laboratory data [32]. In challenging cases, Fibresolve was able to classify those that did not meet the criteria for definite/probable UIP as IPF, with an estimated sensitivity of 56–65% and specificity of 92–94% [33].

Using radiologist-determined visual UIP as baseline, a fully automated supervised CNN was developed to predict the likelihood of UIP on volumetric chest CT scan using a linear support vector machine [34]. The study included 1934 patients for training, 408 for validation and 565 for test performance. The DL-based UIP classifier demonstrated strong performance, achieving a sensitivity of 93% and a specificity of 86%. When applied to a multicentre ILD-only cohort with higher UIP prevalence, the positive predictive value increased to 85%, while the negative predictive value declined to 70%.

Handcrafted radiomics (HCR) is a quantitative method that analyses and extracts high-dimensional imaging features. The potential of HCR ensembled with supervised DL was examined to differentiate between IPF and non-IPF ILD patients on HRCT scans [35]. The study included 474 HRCT scans and a five-fold cross-validation on 365 HRCT scans. The ensemble of HCR and DL models demonstrated the best accuracy and was better than the HCR or DL models alone, indicating that these approaches can complement each other for IPF diagnosis. Moreover, the performance of the proposed models was better than the diagnostic performance of two radiologists and one pulmonologist using an *in silico* trial [35].

AI-driven approaches to prognostic assessment in IPF

Imaging combined with DL has been also proposed to stratify patients based on disease progression, particularly for predicting UIP and evaluating fibrosis severity. Thus, a supervised DL and CNN was developed to predict definite or probable UIP by analysing virtual lung wedges on HRCT, using histopathological UIP as the reference standard [36]. This model was trained on 221 patients and validated on a hold-out cohort of 80 patients, and demonstrated moderate accuracy but was significantly associated with transplant-free survival. Similarly, SOFIA (Systematic Objective Fibrotic Imaging Analysis Algorithm), based on unsupervised DL, was designed to quantify UIP-like features on HRCT in a large cohort of IPF patients [37]. SOFIA was trained on a database of 1157 fibrotic lung disease-specific HRCT scans and validated against the performance of 92 thoracic radiologists on a test cohort of 150 HRCT scans from an independent institution. Unlike human interpretation, this unsupervised DL CNN eliminates perceptual bias by providing continuous UIP probability scores, outperforming expert radiologists and guideline-based histological patterns in predicting patient outcomes.

More recently, a new version of Fibresolve, a vision transformer-based model integrating CT, pulmonary function, age and sex, predicted mortality in ILD across two cohorts (n=220 and n=407) with hazard ratios up to 5.82. Performance remained robust at follow-up and with score changes over time [38].

Another supervised DL model quantified fibrosis by calculating the proportion of honeycombing in the lung and developing a CT fibrosis score (CTS). Patients (n=16 in the training set and n=41 in the verification set) were classified into three fibrosis severity stages (CTS <5, 5–25 and >25), and incorporated into a pulmonary fibrosis severity model (CTPF model) that also included pulmonary function, age and gender. The CTPF model proved to be a reliable predictor of mortality risk over 1, 2 and 3 years, although only marginally superior to CTS or PFT-based assessments alone [39].

Supervised DL-driven CT fibrosis quantification was also used to examine the proportions of normal and fibrotic lung volume and their impact in prognosis [40]. The study, that included 161 patients using commercial DL software, demonstrated that CT-quantified volumetric parameters correlated with forced vital capacity (FVC) and D_{LCO}, and served as independent predictors of overall survival when adjusted for clinical and physiological variables.

In efforts to achieve automatic histological UIP prediction from HRCT, a supervised DL model incorporating multiple instance learning (MIL) for binary classification of UIP was developed and validated across 2143 CT scans [41]. This MIL-UIP classification demonstrated higher sensitivity than

visual assessment in predicting histological UIP, and was linked to worse survival and greater FVC decline. Notably, the relative impact of UIP classification diminished as fibrosis extent increased, aligning with prior findings that quantitative supervised learning CT-based lung fibrosis extent is a strong predictor of functional decline and survival [42].

Moreover, a single meta-cohort analysis suggests that MIL-UIP may uncover genotype–phenotype relationships, as multivariate analysis revealed significant associations between MIL-UIP scores and variants in *MUC5B* (mucin 5B) and *ZKSCAN1* (zinc finger with KRAB and SCAN domains 1), highlighting a putative link between these genes and computational UIP patterns [43].

CALIPER (Computer-Aided Lung Informatics for Pathology Evaluation and Ratings) is an image analysis tool developed by the Biomedical Imaging Resource Laboratory at the Mayo Clinic Rochester (Rochester, MN, USA) for the characterisation and quantification of lung parenchymal findings on HRCT that applies a supervised classifier to label lung voxels as normal, ground-glass, reticulation or honeycombing [44]. The quantification of volume of reticular densities and total volume of interstitial abnormalities has been shown to be predictive of survival after a median follow-up of 2.4 years [44]. Also, pulmonary vessel volume and honeycombing scored by CALIPER were independent predictive variables of mortality [45].

Even quantitative imaging obtained with low-dose CT and processed with CALIPER, independently validated across multiple cohorts, has shown that the UIP pattern identified by supervised algorithms carries a several-fold increased risk of all-cause mortality [46].

DL-based segmentation models have also been applied to quantify lung, airway, vascular and fibrosis volumes from baseline HRCT scans. Using a supervised approach in a cohort of 446 patients (n=223 in the derivation set and n=223 in the validation set), the model demonstrated that reduced lung volume and increased vascular, airway and fibrosis volumes were independently associated with lower 5-year survival, even after adjustment for baseline disease severity [47]. In longitudinal analyses, progressive decline in lung volume on follow-up was significantly linked to higher mortality risk.

Lung fibrosis quantification using data-driven texture analysis (DTA), a CNN-based supervised DL technique, was tested in 393 patients, showing a strong correlation with FVC and D_{LCO} decline and an increased risk of mortality [48].

Multimodal integration

Several multimodal models have also been proposed to enhance diagnostic accuracy and predict prognosis. Thus, a multimodal supervised AI algorithm was developed combining CT imaging, including 28 distinct features associated with the presence/absence of UIP, with a previously described histopathological module [21]. This model was evaluated in a cohort of 114 patients and, when tested on an independent dataset, achieved an AUC of 0.92 for UIP detection. Moreover, training with the algorithm significantly enhanced diagnostic concordance with expert pathologists and boosted diagnostic confidence among general pathologists in identifying UIP [49].

Likewise, to classify different types of ILD, Mei *et al*. [50] developed a CNN model integrating CT images with clinical, functional and morphological data. The performance of this supervised learning AI model was compared to seven human readers, including an expert thoracic radiologist and an expert pulmonologist. The model demonstrated modest performance, with varying results across different ILD types, showing higher sensitivity but lower specificity in some cases, and vice versa. Overall, its performance was comparable to that of human readers [50].

Fibrosis-Net, a supervised DL CNN model, was designed to predict continuous clinical outcomes, by integrating CT scans with spirometry measurements and clinical metadata [51]. Fibrosis-Net estimates future FVC at specific time-points, and both quantitative and qualitative evaluations demonstrated that the model achieves high predictive accuracy. Of interest, Fibrosis-Net is publicly available through the OpenMedAI initiative (https://github.com/darwinai/FibrosisNet).

Likewise, Fibro-CoSANet, a multimodal supervised learning-based approach, was also developed to predict the slope of FVC in an end-to-end manner. Fibro-CoSANet utilised CT images, demographic information (sex, age and smoking) and the initial FVC, in CNN frameworks with a stacked attention layer [52]. Although with interesting results, a major limitation of this approach was that the prognosis of pulmonary fibrosis was of a linear nature, which makes it impossible to predict the actual FVC values at each temporal point.

Strengths, limitations and opportunities

Overall, these HRCT-based AI studies illustrate the breadth of methodological approaches being tested in IPF and ILD, from supervised CNNs and vision transformers trained to replicate radiologist expertise, to multimodal models integrating clinical and functional data, and more recent unsupervised frameworks capable of simulating disease trajectories. Supervised models demonstrate strong accuracy and reproducibility, particularly for distinguishing UIP from non-UIP, and offer the advantage of interpretability by aligning with radiological patterns. However, they remain constrained by reliance on labelled data and expert definitions, which may limit discovery of novel imaging signatures. In contrast, unsupervised methods show promise for uncovering latent disease features and modelling progression, but these approaches are still in an early state, with limited validation. Across all studies, common limitations include small or single-centre cohorts, heterogeneity of imaging protocols, lack of independent external validation and uncertainty about generalisability across diverse populations. A further challenge is interpretability because while deep models may outperform radiologists, their "black box" nature raises concerns for clinical adoption.

Also, the performance of numerous AI algorithms declines when exposed to discrepancies between their training data and real-world environments. This phenomenon, known as the "dataset shift problem", which often remains ignored during model training, raises concerns about the reliability of AI in clinical practice [53].

Actually, preliminary results from a recent study revealed that DL models for IPF diagnosis lack robustness when applied to CT series acquired with different imaging protocols [54]. The authors used a generalised linear mixed effects model to identify key factors contributing to discrepancies in model performance. Their analysis revealed that when three high-performing IPF diagnosis models were applied to CT series collected under different imaging protocols, all three models exhibited a decrease in specificity. This highlights the importance of carefully considering acquisition and reconstruction conditions when designing and implementing DL models in clinical settings [54].

Moreover, AI models trained and validated on existing imaging datasets may show reduced performance when applied to scans from newer CT scanners, particularly those that use proprietary technologies, which modify image characteristics such as resolution, contrast or reconstruction algorithms, differing from legacy systems.

Moving forward, standardisation of imaging acquisition, development of large multicentre datasets and prospective validation are essential. Equally important will be the integration of imaging AI with genomic, proteomic and clinical data to enable robust, clinically meaningful biomarkers of IPF diagnosis and prognosis.

Hyperpolarized xenon-129 magnetic resonance imaging

Hyperpolarized xenon-129 magnetic resonance imaging (HP ^{129}Xe MRI) was developed to simultaneously assess multiple pulmonary microcompartments, providing non-invasive, region-specific measurements of gas transfer at the alveolar–capillary level, including interstitial barrier thickness and red blood cell (RBC) transfer efficiency [55, 56]. CNNs and recurrent neural networks have improved static and dynamic MRI reconstruction [57, 58]. Four distinct gas exchange patterns, often coexisting within the same patient, have been described and align with the heterogeneous histopathology of IPF, where fibrotic areas, honeycombing and normal lung regions can coexist [56]. Both barrier uptake and RBC transfer significantly correlate with D_{LCO} [56]. IPF patients who progressed over 1 year showed reduced RBC-to-barrier ratios and high ventilation percent at baseline compared to non-progressors, despite both groups showing lower gas exchange in fibrotic areas [59]. Preliminary findings also suggest that a regional HP ^{129}Xe MRI biomarker of gas exchange improved with antifibrotic therapy, even when D_{LCO} and other PFTs remained unchanged, although these results are limited by small sample size and baseline differences between treatment groups [60]. Recently, a CNN-based pipeline for HP ^{129}Xe ventilation image registration and segmentation was developed using data from 205 participants and validated in an external cohort of 71 participants. The CNN-derived ventilation defect percent showed strong agreement with expert-derived estimates using a well-established semi-automated approach across three independent sites [61].

Blood biomarkers

Using three publicly available blood transcriptomic datasets and a ML approach, three distinct patient clusters in IPF were identified, each with significant differences in lung function and survival [62]. This method involved both unsupervised (co-normalisation, pooling and clustering of blood transcriptomic data)

and supervised learning components (a gene-based classifier developed and validated on independent datasets). Results showed that Cluster 1 was primarily associated with mitochondrial dysfunction, Cluster 2 with apoptosis and cell cycle pathways, and Cluster 3 with genes related to the immune response. Notably, patients in Cluster 2 demonstrated markedly better survival outcomes.

In a more recent unsupervised consensus clustering ML study, three blood biomarker profiles linked to varying survival rates and functional decline were also revealed [63]. These clusters were classified as the basement membrane cluster, characterised by high levels of basement membrane-associated extracellular matrix (ECM) biomarkers; the epithelial injury cluster, with elevated concentrations of epithelial dysfunctional markers such as matrix metalloproteinase 7, surfactant protein D and cytokeratin 19 fragment; and the cross-linked fibrin cluster, marked by high cross-linked fibrin levels. Among these, the epithelial injury cluster exhibited the highest risk of mortality and the most significant decline in lung function, supporting its recognised pathogenic feature in IPF, whereas the basement membrane cluster had the best survival [63].

Together, these findings reinforce the idea that IPF consists of multiple distinct endotypes, each with varying trajectories of functional decline and survival outcomes.

Recently, using a combination of ML models applied to high-throughput proteomics data from circulating plasma, a proteomic classifier capable of distinguishing cases of connective tissue disease-associated ILD (CTD-ILD) from IPF was developed and validated [64]. This supervised learning ML-derived proteomic classification model exhibited high discriminatory power, with Harrell's C-statistic values ranging from 0.84 to 0.95. The results indicated that a blood-based protein classifier incorporating 37 proteins, sex and age helps to better characterise protein differences between CTD-ILDs and IPF. Moreover, the proteins identified in this study aligned with the hypothesis that CTD-ILD would associate with immune response proteins, whereas IPF would be characterised by proteins associated with epithelial cells, fibrosis and ECM turnover [64].

Transcriptomic signatures

ML and gene expression profiling have been extensively utilised to develop diagnostic and prognostic models for IPF.

For diagnostic purposes, the transcriptomic signatures obtained from three Gene Expression Omnibus datasets, GSE24206, GSE10667 and GSE32537, were used to identify differentially expressed genes between IPF and normal lungs (unsupervised clustering technique). Two supervised ML algorithms, LASSO (Least Absolute Shrinkage and Selection Operator) regression and a random forest (RF) classifier, were used to identify a panel of diagnostic biomarkers for IPF. The LASSO regression algorithm identified five and the RF classifier identified 11 genes with diagnostic potential. The intersection of the two methods revealed three diagnostic biomarkers: *ASPN* (asporin), an ECM component that play an important role in tissue injury and regeneration; *COMP* (cartilage oligomeric matrix protein), a member of the thrombospondin gene family; and *GPX8* (glutathione peroxidase 8). The AUCs for *ASPN*, *COMP* and *GPX8* were 0.94, 0.99 and 0.94, respectively, indicating remarkable diagnostic efficacy for all three genes. Interestingly, single-cell transcriptomic data showed a prominent expression in fibroblasts [65].

IPF is an ageing-associated disease, although the underlying mechanisms remain poorly defined [66]. Recently, the parallels between ageing and IPF were examined using Precious3GPT, an advanced multimodal AI platform trained to generate transcriptomic lung ageing and IPF clocks [67]. The analysis identified only 15 overlapping genes, representing core processes common to both conditions. These findings suggest that while ageing increases susceptibility to IPF, the disease is driven by distinct pathological pathways that diverge substantially from normal age-related changes.

On the other hand, to construct predictive models for disease progression and survival, several studies have used the GSE70866 dataset, which contains gene expression data from bronchoalveolar lavage samples from 176 IPF patients.

In one of the studies, an ECM-related prognostic model based on the expression of seven genes (*CTS6* (cystatin-E/M), *PPBP* (pro-platelet basic protein), *CSPG4* (chondroitin sulfate proteoglycan 4), *SEMA3B* (semaphorin 3B), *LAMB2* (laminin subunit beta-2), *SERPINB4* (serpin family B member 4) and *CTF1* (cardiotrophin 1)) was developed. This supervised model successfully stratified patients into high- and low-risk groups, with high-risk patients exhibiting significantly lower survival rates [68]. The model's

reliability was further validated using an external dataset (GSE28042), whose gene expression data were obtained from peripheral blood mononuclear cells.

A subsequent study applied RSF (Random Survival Forest) modelling to the same GSE70866 dataset, integrating transcriptional and clinical data [69]. This ML-based approach (supervised learning, integrated with unsupervised techniques for data pre-processing and feature extraction) identified four prognostic genes (*DTL* (denticleless E3), *CEP55* (centrosomal protein 55), *PCLAF* (PCNA clamp associated factor) and *KIF23* (kinesin family member 23)). The derived risk score effectively stratified patients into high- and low-risk groups, with the high-risk group showing significant enrichment in ECM receptor interaction, mitogen-activated protein kinase signalling and focal adhesion pathways [69].

In another study, researchers developed an immune-related risk score (IRS) prognostic model using an unsupervised ML integration framework [70]. This model was also constructed using GSE70866, along with two additional datasets (GSE110147 and GSE10667), which were generated using tissue microarrays on different platforms (Agilent Technologies and Affymetrix) and few patients. By applying 10 different ML algorithms, researchers identified a 20-gene prognostic signature, which was used to calculate the IRS for each patient. Survival analysis revealed that patients in the high-IRS group had significantly worse survival than those in the low-IRS group [70]. Additionally, analysis of GSE70866 found that high *PLA2G7* (phospholipase A2 group VII) expression was linked to a significant reduction in overall survival.

Therefore, despite using similar datasets, primarily GSE70866, these studies yielded distinct findings, highlighting challenges in comparing diverse biological samples, the complexity of IPF, the variability of ML methodologies and the use of small cohorts.

An additional problem is the so-called “black box” that represents the difficulty of understanding how AI models make decisions, *i.e.* what information these methods use when making predictions. This lack of transparency can make it hard to fix errors and improve the models. In this context, a biologically informed DL model, structured around biological pathway ontologies, shows a promising strategy to mitigate the “black box” nature of multi-omics data analysis [71].

Challenges and future directions for AI integration in IPF

While all these computational models highlight the potential of AI-driven prognostic tools in IPF, large-scale, prospective, multicentre studies with rigorous clinical validation are needed to confirm their clinical applicability.

Ideally, a multimodal DL model that integrates clinical information, HRCT imaging, pathological features and multi-omics data could significantly improve the accurate prediction of IPF risk progression (figure 3). This approach has been recently proposed in non-small cell lung cancer, demonstrating enhanced survival prediction even when using relatively small datasets [72].

Nevertheless, despite promising performance, integrating AI into IPF clinical management faces substantial challenges. High costs, complex implementation into hospital systems, especially in resource-limited settings, and the prevalence of proprietary, non-transparent algorithms trained on distinct datasets make it difficult for clinicians and institutions to identify the most reliable or generalisable models [73]. Real-world distribution requires unified integration with digital platforms, clinician-friendly interfaces and robust validation across diverse populations.

Although large datasets can reduce variability they do not guarantee representativeness; many models rely on data from expert centres that under-represent patients from lower socioeconomic or minority populations, raising ethical concerns and risking poor performance in these populations [74]. Regional and genetic differences in IPF pathogenesis may further limit generalisability. Overcoming these challenges demands technical solutions and structural reforms in data sharing, inclusion and regulatory oversight. Transparent performance evaluation, large collaborative validations and supportive policy frameworks are essential for safe, equitable AI adoption in IPF care as promoted through reporting guidelines [75]. Table 2 summarises the key questions and uncertainties in this field.

The contribution of AI in drug discovery and innovation

Traditional drug development faces multiple obstacles, such as extended timelines, high costs and significant failure rates. AI has the potential to expedite the transition of candidate drugs into clinical practice, improving the efficiency and cost-effectiveness.

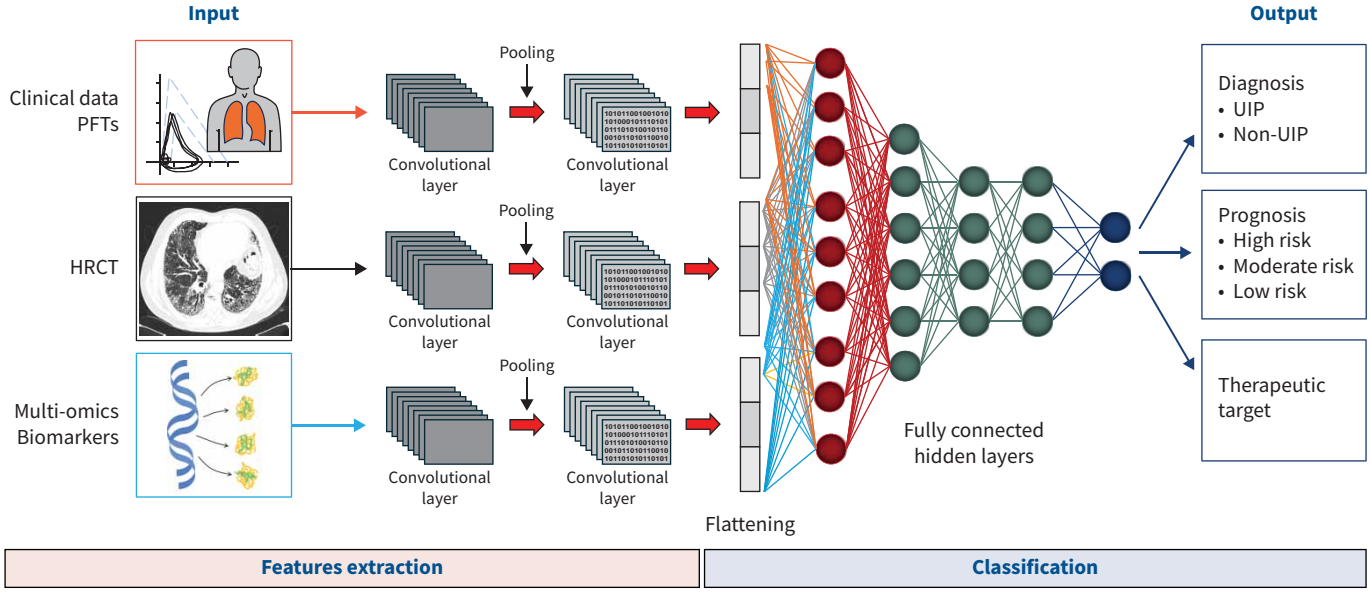


FIGURE 3 Integrated deep learning framework for multimodal diagnosis, prognosis and identification of therapeutic targets in usual interstitial pneumonia (UIP). The diagram illustrates an advanced convolutional neural network-based model designed to integrate heterogeneous data sources for comprehensive assessment of interstitial lung disease. The framework simultaneously processes three primary data modalities: 1) clinical data, including demographic information, pulmonary function test (PFT) results, comorbidities and blood-based biomarkers; 2) high-resolution computed tomography (HRCT) images capturing detailed lung parenchymal features; and 3) multi-omics data, including genomic, transcriptomic and proteomic profiles. Each modality is independently passed through multiple convolutional and pooling layers to extract modality-specific representations. These features are then flattened and combined in fully connected hidden layers, allowing the model to learn complex cross-modal interactions. The integrated output layer provides multiple predictions, including diagnostic classification (UIP *versus* non-UIP), prognostic risk stratification (high, moderate or low risk) and discovery of potential therapeutic targets, thereby enhancing precision diagnosis and personalised clinical decision making.

Several AI techniques have been applied for drug discovery, including target identification, virtual screening, and prediction of absorption, distribution, metabolism, excretion and toxicity [76, 77]. A recent analysis performed in the global Pharmaprojects research and development database (www.citeline.com/en/

TABLE 2 Key questions and areas of uncertainty in the application artificial intelligence (AI) in idiopathic pulmonary fibrosis (IPF)

Key question/uncertainty	Description	Implications
Generalisability of models	Can AI models trained on specific cohorts maintain performance across genetically and environmentally diverse populations?	Requires multicentre validation and external testing with fair representation of diverse patient groups
Interpretability and transparency	Can clinicians understand the rationale behind AI-generated decisions?	Influences clinical confidence, adoption and regulatory approval
Integration of multimodal data	What are the optimal strategies to combine imaging, clinical and molecular/genomic datasets?	Requires standardisation and alignment across data types
Disease heterogeneity	Can AI reliable differentiate meaningful IPF subtypes or molecular endotypes?	Crucial for personalised prognosis and targeted therapies
Reference label precision	Are current diagnostic labels (*e.g.* UIP/IPF) accurate and consistent enough for supervised training?	Directly impacts model accuracy, reproducibility and comparability
Longitudinal prediction	Can AI effectively anticipate disease progression at the individual level?	Critical for patient counselling, therapeutic planning and trial design
Ethical and regulatory considerations	How should AI systems address data privacy, patient consent and bias mitigation?	Determines safe, fair and standardised clinical implementation
Clinical relevance and adoption	Do AI tools lead to measurable improvements in patient outcomes when implemented in practice?	Requires prospective evaluation and real-world validation

UIP: usual interstitial pneumonia.

products-services/clinical/pharmaprojects) revealed that AI has been used in the development of 164 investigational drugs [78].

Furthermore, AI may expand the available drug pool by repurposing existing approved medications using large-scale biomedical datasets [77, 79].

Recently, researchers utilised PandaOmics, a commercially available target discovery platform (https://pharma.ai/pandaomics), which integrates multi-omics datasets from IPF lungs, biological network analysis and text mining from scientific literature [79]. Through this computational approach, TRAF2- and NCK-interacting kinase (TNIK) was identified as a potential target for antifibrotic therapy [80]. To discover TNIK inhibitors, the available crystal structures of the TNIK kinase domain [81] were employed in conjunction with the Chemistry42 structure-based drug design AI workflow. This approach led to the development of a specific TNIK inhibitor, INS018_055, that was shown to mitigate transforming growth factor (TGF)-β-induced epithelial-to-mesenchymal transition and fibroblast-to-myofibroblast differentiation, and also to reduce bleomycin-induced lung fibrosis in mice [80]. Additionally, two phase 1 studies confirmed that INS018_055 is safe, well tolerated and exhibits good oral bioavailability with dose-proportional pharmacokinetics [80].

Finally, the phase 2a results for this inhibitor, now known as rentosertib, were recently published [82]. At 12 weeks, the 60 mg once-daily dose produced a mean FVC increase of 98.4 (95% CI 10.9–185.9) mL compared with a mean change of −20.3 (95% CI −116.1–75.6) mL in the placebo group. The study also confirmed the drug's safety, good tolerability and favourable pharmacokinetic profile. Likewise, employing an integrated strategy combining ML-based prediction and structure-based virtual screening to repurpose drugs for IPF treatment, it was found that osimertinib, an epidermal growth factor receptor tyrosine kinase inhibitor, has potent TNIK inhibitory activity [83].

Interestingly, beyond its antifibrotic effects, TNIK inhibition also reduced cellular senescence and a senescence-associated secretory phenotype in various experimental senescence models [84], a process that plays a critical profibrotic role in IPF [85].

Combining unsupervised network analysis and DL-based small-molecule compound screening frameworks for therapeutic drug discovery, several novel candidate genes and a number of small-molecule compounds potentially useful in IPF were identified [86]. Among them, a glutaminase 1 inhibitor (CB-839), which has been shown to be therapeutically efficacious in treating both bleomycin- and TGF-β1-induced pulmonary fibrosis in mice, was revealed [87].

A recent report in 2024 introduced Standigm ASK, an innovative AI-assisted drug discovery platform which combines a heterogeneous knowledge graph database with a neural network to uncover novel therapeutic targets [88]. Standigm ASK operates based on five key strategic criteria: biological relevance, disease causality, suitability, toxicity and novelty. Using this approach, researchers initially identified several genes potentially involved in epithelial-to-mesenchymal transition in IPF.

Finally, DeepMind's AlphaFold2, a DL AI system, has revolutionised drug discovery by predicting protein structures with near-experimental accuracy, accelerating target identification and structure-based design [89]. The newer AlphaFold3 further advances these capabilities, although experimental validation remains essential [90].

Conclusions

AI has demonstrated appreciable performance across a variety of clinical tasks, including diagnosis, prediction of survival and treatment.

In IPF, AI-powered imaging (HRCT) analysis has improved the early and accurate diagnosis of this disease. These models can identify subtle fibrosis patterns that radiologists might miss, enabling earlier intervention and improving patient outcomes. Beyond diagnosis, ML/DL models trained on datasets of HRCT, clinical and genomic information can predict disease progression. Very recently, AI has also strongly contributed in drug discovery by identifying a novel and promising therapeutic target.

However, implementing AI models in the diagnosis and management of IPF presents several challenges. One significant issue is the need for large, high-quality datasets, from different institutions, to train these models effectively. IPF is a relatively infrequent disease, which limits the availability of extensive, well-annotated datasets needed for DL algorithms.

Additionally, in the case of HRCT where most studies have been performed, variations in imaging protocols, scanner types and patient populations across different medical institutions can introduce inconsistencies, potentially reducing the generalisability of AI models. Also, it is important to consider that IPF is a highly heterogeneous disease, with patients showing diverse disease trajectories, transcriptomic profiles, and lung and blood biomarkers, and most AI models use static snapshots missing temporal patterns.

Another major challenge that faces any medical purposing is the interpretability and reliability of AI-driven decisions. Many DL models operate as "black boxes", meaning that we cannot fully explain their decision-making processes or how the model arrived at its output. Ensuring that AI models are rigorously validated, explainable and impeccably integrated into existing clinical workflows is essential for maximising their impact in IPF management.

Also important is the potential for AI to exacerbate existing health disparities, especially when access to these tools is restricted in resource-limited settings.

Despite these challenges and unforeseen problems, AI offers substantial potential to enhance patient-centred diagnostics and outcome assessments.

Ultimately, we hope that AI will serve as a tool to enhance rather than replace human creativity, because the true spark of imagination, emotion and personal experience, the essence of creativity, remains deeply human.

Acknowledgements: The authors appreciate the important collaboration of Yalbi Balderas Martínez (INER "Ismael Cosío Villegas", Mexico City, Mexico) in the design of the figures.

Author contributions: Concept and design: M. Selman. Drafting of manuscript, critical revision and approval of the final draft of the paper: all authors

Conflict of interest: The authors have no potential conflicts of interest to declare.

References

1 Raghu G, Remy-Jardin M, Richeldi L, *et al.* Idiopathic pulmonary fibrosis (an update) and progressive pulmonary fibrosis in adults: an official ATS/ERS/JRS/ALAT clinical practice guideline. *Am J Respir Crit Care Med* 2022; 205: e18–e47.
2 Peng Y, Zhang Y, Wang L. Artificial intelligence in biomedical engineering and informatics: an introduction and review. *Artif Intell Med* 2010; 48: 71–73.
3 Haug CJ, Drazen JM. Artificial intelligence and machine learning in clinical medicine. *N Engl J Med* 2023; 388: 1201–1208.
4 Rowe M. An introduction to machine learning for clinicians. *Acad Med* 2019; 94: 1433–1436.
5 Sarker IH. Deep learning: a comprehensive overview on techniques, taxonomy, applications and research directions. *SN Comput Sci* 2021; 2: 420.
6 Wang TW, Tzeng YH, Wu KT, *et al.* Meta-analysis of deep learning approaches for automated coronary artery calcium scoring: performance and clinical utility AI in CAC scoring: a meta-analysis: AI in CAC scoring: a meta-analysis. *Comput Biol Med* 2024; 183: 109295.
7 Greener JG, Kandathil SM, Moffat L, *et al.* A guide to machine learning for biologists. *Nat Rev Mol Cell Biol* 2022; 23: 40–55.
8 Galic I, Habijan M, Leventic H, *et al.* Machine learning empowering personalized medicine: a comprehensive review of medical image analysis methods. *Electronics* 2023; 12: 4411.
9 Richeldi L, Cottin V, Würtemberger G, *et al.* Digital lung auscultation: will early diagnosis of fibrotic interstitial lung disease become a reality? *Am J Respir Crit Care Med* 2019; 200: 261–263.
10 Moran-Mendoza O, Ritchie T, Aldhaheri S. Fine crackles on chest auscultation in the early diagnosis of idiopathic pulmonary fibrosis: a prospective cohort study. *BMJ Open Respir Res* 2021; 8: e000815.
11 Sethi AK, Muddaloor P, Anvekar P, *et al.* Digital pulmonology practice with phonopulmography leveraging artificial intelligence: future perspectives using dual microwave acoustic sensing and imaging. *Sensors* 2023; 23: 5514.
12 Horimasu Y, Ohshimo S, Yamaguchi K, *et al.* A machine-learning based approach to quantify fine crackles in the diagnosis of interstitial pneumonia: a proof-of-concept study. *Medicine* 2021; 100: e24738.
13 Gompelmann D, Gysan MR, Desbordes P, *et al.* AI-powered evaluation of lung function for diagnosis of interstitial lung disease. *Thorax* 2025; 80: 445–450.

14 Kim H, Jin KN, Yoo SJ, *et al.* Deep learning for estimating lung capacity on chest radiographs predicts survival in idiopathic pulmonary fibrosis. *Radiology* 2023; 306: e220292.
15 Smith ML. The histologic diagnosis of usual interstitial pneumonia of idiopathic pulmonary fibrosis. Where we are and where we need to go. *Mod Pathol* 2022; 35: Suppl. 1, 8–14.
16 Kim SY, Diggans J, Pankratz D, *et al.* Classification of usual interstitial pneumonia in patients with interstitial lung disease: assessment of a machine learning approach using high-dimensional transcriptional data. *Lancet Respir Med* 2015; 3: 473–482.
17 Pankratz DG, Choi Y, Imtiaz U, *et al.* Usual interstitial pneumonia can be detected in transbronchial biopsies using machine learning. *Ann Am Thorac Soc* 2017; 14: 1646–1654.
18 Choi Y, Liu TT, Pankratz DG, *et al.* Identification of usual interstitial pneumonia pattern using RNA-Seq and machine learning: challenges and solutions. *BMC Genomics* 2018; 19: Suppl. 2, 101.
19 Raghu G, Flaherty KR, Lederer DJ, *et al.* Use of a molecular classifier to identify usual interstitial pneumonia in conventional transbronchial lung biopsy samples: a prospective validation study. *Lancet Respir Med* 2019; 7: 487–496.
20 Hou J, Yang Y, Han X. Machine learning and single-cell analysis identify molecular features of IPF-associated fibroblast subtypes and their implications on IPF prognosis. *Int J Mol Sci* 2023; 25: 94.
21 Uegami W, Bychkov A, Ozasa M, *et al.* MIXTURE of human expertise and deep learning-developing an explainable model for predicting pathological diagnosis and survival in patients with interstitial lung disease. *Mod Pathol* 2022; 35: 1083–1091.
22 Mäkelä K, Mäyränpää MI, Sihvo HK, *et al.* Artificial intelligence identifies inflammation and confirms fibroblast foci as prognostic tissue biomarkers in idiopathic pulmonary fibrosis. *Hum Pathol* 2021; 107: 58–68.
23 Tiitto L, Bloigu R, Heiskanen U, *et al.* Relationship between histopathological features and the course of idiopathic pulmonary fibrosis/usual interstitial pneumonia. *Thorax* 2006; 61: 1091–1095.
24 King TE Jr, Schwarz MI, Brown K, *et al.* Idiopathic pulmonary fibrosis: relationship between histopathologic features and mortality. *Am J Respir Crit Care Med* 2001; 164: 1025–1032.
25 Harada T, Watanabe K, Nabeshima K, *et al.* Prognostic significance of fibroblastic foci in usual interstitial pneumonia and non-specific interstitial pneumonia. *Respirology* 2013; 18: 278–283.
26 Walsh SLF, Calandriello L, Silva M, *et al.* Deep learning for classifying fibrotic lung disease on high-resolution computed tomography: a case-cohort study. *Lancet Respir Med* 2018; 6: 837–845.
27 Bratt A, Williams JM, Liu G, *et al.* Predicting usual interstitial pneumonia histopathology from chest CT imaging with deep learning. *Chest* 2022; 162: 815–823.
28 Choe J, Hwang HJ, Seo JB, *et al.* Content-based image retrieval by using deep learning for interstitial lung disease diagnosis with chest CT. *Radiology* 2022; 302: 187–197.
29 Xu MW, Zhang ZH, Wang X, *et al.* PMFF-Net: a deep learning-based image classification model for UIP, NSIP, and OP. *Comput Biol Med* 2025; 195: 110618.
30 Maddali MV, Kalra A, Muelly M, *et al.* Development and validation of a CT-based deep learning algorithm to augment non-invasive diagnosis of idiopathic pulmonary fibrosis. *Respir Med* 2023; 219: 107428.
31 Bradley J, Huang J, Kalra A, *et al.* External validation of Fibresolve, a machine-learning algorithm, to non-invasively diagnose idiopathic pulmonary fibrosis. *Am J Med Sci* 2024; 367: 195–200.
32 Ahmad Y, Mooney J, Allen IE, *et al.* A machine learning system to indicate diagnosis of idiopathic pulmonary fibrosis non-invasively in challenging cases. *Diagnostics* 2024; 14: 830.
33 Chang M, Reicher JJ, Kalra A, *et al.* Analysis of validation performance of a machine learning classifier in interstitial lung disease cases without definite or probable usual interstitial pneumonia pattern on CT using clinical and pathology-supported diagnostic labels. *J Imaging Inform Med* 2024; 37: 297–307.
34 Chung JH, Chelala L, Pugashetti JV, *et al.* A deep learning-based radiomic classifier for usual interstitial pneumonia. *Chest* 2024; 165: 371–380.
35 Refaee T, Salahuddin Z, Frix AN, *et al.* Diagnosis of idiopathic pulmonary fibrosis in high-resolution computed tomography scans using a combination of handcrafted radiomics and deep learning. *Front Med* 2022; 9: 915243.
36 Shaish H, Ahmed FS, Lederer D, *et al.* Deep learning of computed tomography virtual wedge resection for prediction of histologic usual interstitial pneumonitis. *Ann Am Thorac Soc* 2021; 18: 51–59.
37 Walsh SLF, Mackintosh JA, Calandriello L, *et al.* Deep learning-based outcome prediction in progressive fibrotic lung disease using high-resolution computed tomography. *Am J Respir Crit Care Med* 2022; 206: 883–891.
38 Callahan SJ, Scholand MB, Kalra A, *et al.* Multi-modal machine learning classifier for idiopathic pulmonary fibrosis predicts mortality in interstitial lung diseases. *Respir Investig* 2025; 63: 1012–1017.
39 Wu X, Yin C, Chen X, *et al.* Idiopathic pulmonary fibrosis mortality risk prediction based on artificial intelligence: the CTPF model. *Front Pharmacol* 2022; 13: 878764.
40 Nam JG, Choi Y, Lee SM, *et al.* Prognostic value of deep learning-based fibrosis quantification on chest CT in idiopathic pulmonary fibrosis. *Eur Radiol* 2023; 33: 3144–3155.

41 Humphries SM, Thieke D, Baraghoshi D, *et al.* Deep learning classification of usual interstitial pneumonia predicts outcomes. *Am J Respir Crit Care Med* 2024; 209: 1121–1131.
42 Oh AS, Lynch DA, Swigris JJ, *et al.* Deep learning-based fibrosis extent on computed tomography predicts outcome of fibrosing interstitial lung disease independent of visually assessed computed tomography pattern. *Ann Am Thorac Soc* 2024; 21: 218–227.
43 Blumhagen RZ, Humphries SM, Peljto AL, *et al. MUC5B* genotype and other common variants are associated with computational imaging features of usual interstitial pneumonia. *Ann Am Thorac Soc* 2025; 22: 533–540.
44 Maldonado F, Moua T, Rajagopalan S, *et al.* Automated quantification of radiological patterns predicts survival in idiopathic pulmonary fibrosis. *Eur Respir J* 2014; 43: 204–212.
45 Jacob J, Bartholmai BJ, Rajagopalan S, *et al.* Mortality prediction in idiopathic pulmonary fibrosis: evaluation of computer-based CT analysis with conventional severity measures. *Eur Respir J* 2017; 49: 1601011.
46 Wang JM, Bose S, Murray S, *et al.* Quantitative CT measures of lung fibrosis and outcomes in the national lung screening trial. *Ann Am Thorac Soc* 2025; 22: 1314–1320.
47 Thillai M, Oldham JM, Ruggiero A, *et al.* Deep learning-based segmentation of computed tomography scans predicts disease progression and mortality in idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2024; 210: 465–472.
48 Humphries SM, Mackintosh JA, Jo HE, *et al.* Quantitative computed tomography predicts outcomes in idiopathic pulmonary fibrosis. *Respirology* 2022; 27: 1045–1053.
49 Lami K, Ozasa M, Che X, *et al.* Enhancing interstitial lung disease diagnoses through multimodal AI integration of histopathological and CT image data. *Respirology* 2025; 30: 726–735.
50 Mei X, Liu Z, Singh A, *et al.* Interstitial lung disease diagnosis and prognosis using an AI system integrating longitudinal data. *Nat Commun* 2023; 14: 2272.
51 Wong A, Lu J, Dorfman A, *et al.* Fibrosis-net: a tailored deep convolutional neural network design for prediction of pulmonary fibrosis progression from chest CT images. *Front Artif Intell* 2021; 4: 764047.
52 Al Nazi Z, Rabbi Mashrur F, Islam MA, *et al.* Fibro-CoSANet: pulmonary fibrosis prognosis prediction using a convolutional self attention network. *Phys Med Biol* 2021; 66: 225013.
53 Choi Y, Yu W, Nagarajan MB, *et al.* Translating AI to clinical practice: overcoming data shift with explainability. *Radiographics* 2023; 43: e220105.
54 Yu W, McNitt-Gray MF, Goldin JG, *et al.* Evaluating the robustness of deep learning models trained to diagnose idiopathic pulmonary fibrosis using a retrospective study. *Med Phys* 2025; 52: 4239–4249.
55 MacLeod JL, Khan HM, Franklin A, *et al.* Hyperpolarized xenon-129 MRI: narrative review of clinical studies, testing, and implementation of advanced pulmonary *in vivo* imaging and its diagnostic applications. *Diagnostics* 2025; 15: 474.
56 Wang JM, Robertson SH, Wang Z, *et al.* Using hyperpolarized ^{129}Xe MRI to quantify regional gas transfer in idiopathic pulmonary fibrosis. *Thorax* 2018; 73: 21–28.
57 Duan C, Deng H, Xiao S, *et al.* Fast and accurate reconstruction of human lung gas MRI with deep learning. *Magn Reson Med* 2019; 82: 2273–2285.
58 Wu Y, Ma Y, Capaldi DP, *et al.* Incorporating prior knowledge *via* volumetric deep residual network to optimize the reconstruction of sparsely sampled MRI. *Magn Reson Imaging* 2020; 66: 93–103.
59 Hahn AD, Carey KJ, Barton GP, *et al.* Hyperpolarized ^{129}Xe MR spectroscopy in the lung shows 1-year reduced function in idiopathic pulmonary fibrosis. *Radiology* 2022; 305: 688–696.
60 Hahn AD, Carey KJ, Barton GP, *et al.* Functional xenon-129 magnetic resonance imaging response to antifibrotic treatment in idiopathic pulmonary fibrosis. *ERJ Open Res* 2023; 9: 00080-2023.
61 Mozaffaripour A, Matheson AM, Rahman O, *et al.* Pulmonary ^{129}Xe MRI: CNN registration and segmentation to generate ventilation defect percent with multi-center validation. *Acad Radiol* 2025; 32: 1734–1742.
62 Kraven LM, Taylor AR, Molyneaux PL, *et al.* Cluster analysis of transcriptomic datasets to identify endotypes of idiopathic pulmonary fibrosis. *Thorax* 2023; 78: 551–558.
63 Fainberg HP, Moodley Y, Triguero I, *et al.* Cluster analysis of blood biomarkers to identify molecular patterns in pulmonary fibrosis: assessment of a multicentre, prospective, observational cohort with independent validation. *Lancet Respir Med* 2024; 12: 681–692.
64 Huang Y, Ma SF, Oldham JM, *et al.* Machine learning of plasma proteomics classifies diagnosis of interstitial lung disease. *Am J Respir Crit Care Med* 2024; 210: 444–454.
65 Guo Y, Jin Q, Kang Y, *et al.* Integrating machine learning and neural networks for new diagnostic approaches to idiopathic pulmonary fibrosis and immune infiltration research. *PLoS One* 2025; 20: e0320242.
66 Selman M, Pardo A. Revealing the pathogenic and aging-related mechanisms of the enigmatic idiopathic pulmonary fibrosis. An integral model. *Am J Respir Crit Care Med* 2014; 189: 1161–1172.
67 Galkin F, Chen S, Aliper A, *et al.* AI-driven toolset for IPF and aging research associates lung fibrosis with accelerated aging. *Aging* 2025; 17: 1999–2014.
68 Luo H, Yan J, Zhou X. Constructing an extracellular matrix-related prognostic model for idiopathic pulmonary fibrosis based on machine learning. *BMC Pulm Med* 2023; 23: 397.

69 Ruan H, Ren C. Integrating machine learning with bioinformatics for predicting idiopathic pulmonary fibrosis prognosis: developing an individualized clinical prediction tool. *Exp Biol Med* 2024; 249: 10215.

70 Liu T, Ning J, Fan X, *et al.* Identification of immune patterns in idiopathic pulmonary fibrosis patients driven by PLA2G7-positive macrophages using an integrated machine learning survival framework. *Sci Rep* 2024; 14: 22369.

71 Selby DA, Sprang M, Ewald J, *et al.* Beyond the black box with biologically informed neural networks. *Nat Rev Genet* 2025; 26: 371–372.

72 Verma S, Magazzù G, Eftekhari N, *et al.* Cross-attention enables deep learning on limited omics-imaging-clinical data of 130 lung cancer patients. *Cell Rep Methods* 2024; 4: 100817.

73 Weiner EB, Dankwa-Mullan I, Nelson WA, *et al.* Ethical challenges and evolving strategies in the integration of artificial intelligence into clinical practice. *PLoS Digit Health* 2025; 4: e0000810.

74 Chinta SV, Wang Z, Palikhe A, *et al.* AI-driven healthcare: fairness in AI healthcare: a survey. *PLoS Digit Health* 2025; 4: e0000864.

75 Kolbinger FR, Veldhuizen GP, Zhu J, *et al.* Reporting guidelines in medical artificial intelligence: a systematic review and meta-analysis. *Commun Med* 2024; 4: 71.

76 Chakraborty C, Bhattacharya M, Lee SS, *et al.* The changing scenario of drug discovery using AI to deep learning: recent advancement, success stories, collaborations, and challenges. *Mol Ther Nucleic Acids* 2024; 35: 102295.

77 Zhang K, Yang X, Wang Y, *et al.* Artificial intelligence in drug development. *Nat Med* 2025; 31: 45–59.

78 Druedahl LC, Price WN 2nd, Minssen T, *et al.* Use of artificial intelligence in drug development. *JAMA Netw Open* 2024; 7: e2414139.

79 Tanoli Z, Fernández-Torras A, Özcan UO, *et al.* Computational drug repurposing: approaches, evaluation of in silico resources and case studies. *Nat Rev Drug Discov* 2025; 24: 521–542.

80 Ren F, Aliper A, Chen J, *et al.* A small-molecule TNIK inhibitor targets fibrosis in preclinical and clinical models. *Nat Biotechnol* 2025; 43: 63–75.

81 Masuda M, Uno Y, Ohbayashi N, *et al.* TNIK inhibition abrogates colorectal cancer stemness. *Nat Commun* 2016; 7: 12586.

82 Xu Z, Ren F, Wang P, *et al.* A generative AI-discovered TNIK inhibitor for idiopathic pulmonary fibrosis: a randomized phase 2a trial. *Nat Med* 2025; 31: 2602–2610.

83 Zhao L, Liu H, Yao X, *et al.* Integrated machine learning and structure-based virtual screening identify osimertinib as a TNIK inhibitor for idiopathic pulmonary fibrosis. *J Chem Inf Model* 2025; 65: 10673–10687.

84 Tang Q, Xiao D, Veviorskiy A, *et al.* AI-driven robotics laboratory identifies pharmacological TNIK inhibition as a potent senomorphic agent. *Aging Dis* 2026; 17: 432–451.

85 Pardo A, Selman M. The interplay of the genetic architecture, aging, and environmental factors in the pathogenesis of idiopathic pulmonary fibrosis. *Am J Respir Cell Mol Biol* 2021; 64: 163–172.

86 Ghandikota S, Jegga AG. From gene networks to therapeutics: a causal inference and deep learning approach for drug discovery. *Pharmaceuticals* 2025; 18: 1304.

87 Cui H, Xie N, Jiang D, *et al.* Inhibition of glutaminase 1 attenuates experimental pulmonary fibrosis. *Am J Respir Cell Mol Biol* 2019; 61: 492–500.

88 Han S, Lee JE, Kang S, *et al.* Standigm ASK™: knowledge graph and artificial intelligence platform applied to target discovery in idiopathic pulmonary fibrosis. *Brief Bioinform* 2024; 25: bbae035.

89 Borkakoti N, Thornton JM. AlphaFold2 protein structure prediction: implications for drug discovery. *Curr Opin Struct Biol* 2023; 78: 102526.

90 Abramson J, Adler J, Dunger J, *et al.* Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* 2024; 630: 493–500.