Exploring computer-based imaging analysis in interstitial lung disease: opportunities and challenges

Federico N. Felder and Simon L.F. Walsh

National Heart and Lung Institute, Imperial College London, London, UK.

Corresponding author: Federico N. Felder (ffelder@ic.ac.uk)

Shareable abstract (@ERSpublications)
The application of computer image analysis in interstitial lung disease has overcome the limitations of semiquantitative methods and yields more accurate results. However, there are still obstacles to implementation in clinical practice and drug trials. https://bit.ly/3M9H8Nb

Cite this article as: Felder FN, Walsh SLF. Exploring computer-based imaging analysis in interstitial lung disease: opportunities and challenges. *ERJ Open Res* 2023; 9: 00145-2023 [DOI: 10.1183/23120541.00145-2023].

Received: 8 March 2023
Accepted: 3 May 2023

Abstract

The advent of quantitative computed tomography (QCT) and artificial intelligence (AI) using high-resolution computed tomography data has revolutionised the way interstitial diseases are studied. These quantitative methods provide more accurate and precise results compared to prior semiquantitative methods, which were limited by human error such as interobserver disagreement or low reproducibility. The integration of QCT and AI and the development of digital biomarkers has facilitated not only diagnosis but also prognostication and prediction of disease behaviour, not just in idiopathic pulmonary fibrosis in which they were initially studied, but also in other fibrotic lung diseases. These tools provide reproducible, objective prognostic information which may facilitate clinical decision-making. However, despite the benefits of QCT and AI, there are still obstacles that need to be addressed. Important issues include optimal data management, data sharing and maintenance of data privacy. In addition, the development of explainable AI will be essential to develop trust within the medical community and facilitate implementation in routine clinical practice.

Background

Interstitial lung disease (ILD) is a group of disorders characterised by lung tissue inflammation and/or fibrosis. Overall, they represent complex clinical entities with nonspecific pulmonary symptoms and functional findings. Patients present with progressive dyspnoea, dry cough and restrictive patterns on pulmonary function tests. ILD is a broad term that encompasses many different conditions in which inflammation or fibrosis of interstitium is found in variable proportions affecting disease behaviour and response to treatment. At one end of the ILD spectrum is idiopathic pulmonary fibrosis (IPF), a fibrotic disorder with an inexorably progressive course and poor prognosis (3–5 years) [1, 2]. However, there are other ILDs that are mainly characterised by inflammation and have better outcomes with or without treatment, and higher survival rates [3–6]. Although there has been significant progress in treatment of these conditions in the past decade, in an addition to IPF, are other forms of pulmonary fibrosis which progress regardless of treatment and demonstrate an IPF-like disease course. These non-IPF progressive forms of fibrosis have recently been collectively named “progressive pulmonary fibrosis” [7–10]. High-resolution computed tomography (HRCT) of the chest is central to diagnosis in patients with suspected fibrotic lung disease by providing detailed cross-sectional images of lungs and evaluating disease distribution in three dimensions. In addition, HRCT may play a prognostic role in fibrotic lung disease, and given that it is routinely performed in most patients with suspected fibrotic lung disease, is an attractive target for biomarker research in these diseases [7, 11].

At the most basic level, a typical or probable usual interstitial pneumonia (UIP) pattern (so-called UIP-like disease) is associated with a poor prognosis based on recent antifibrotic therapy trials in IPF and progressive non-IPF disease [8, 12–16]. In addition to the HRCT phenotype, specific HRCT patterns can

be visually quantified (known as semiquantitative evaluation) and used as prognostic markers. Honeycombing, a cardinal sign of fibrosis on HRCT and a key pattern in the identification of UIP, is defined as clustered cystic air spaces, cysts of comparable diameters and cyst diameters typically <10 mm surrounded by well-defined walls [17]. When scored for extent visually, either alone or in combination with the extent of reticulation (sometimes called a "fibrosis score"), honeycombing has been linked consistently to mortality idiopathic fibrotic lung disease (IPF and idiopathic nonspecific interstitial pneumonia), connective tissue disease-related fibrotic lung disease (CTD-FLD) and fibrotic hypersensitivity pneumonitis (FHP) over the past two decades [18–23]. In one study involving 315 patients with IPF enrolled in a clinical trial of IFN-γ1b, Lynch *et al*. [23] reported that the overall extent of fibrosis, defined as the extent of reticular and honeycombing patterns combined, was the strongest predictor of mortality. It is noteworthy that in this study, HRCT was a better predictor of mortality than pulmonary function in IPF. The severity of traction bronchiectasis is also a strong predictor of mortality in multiple fibrotic lung disease subsets [19, 20, 22, 24] and may be a sensitive surrogate marker of disease progression in IPF [25]. Most recently, changes in aortosternal distance and fissural displacement measured manually predict outcomes in patients with IPF [26]. In contrast, the presence of certain patterns may be associated with a more favourable outcome. In FHP, the presence of mosaic attenuation and air trapping may be associated with a more favourable survival [27]. Since disease severity based on HRCT fibrosis extent and lung function decline have been reported as independent predictors of outcome, these variables have been combined to create staging systems in IPF, systemic sclerosis related ILD and fibrotic sarcoidosis [23, 28–31].

Despite this large body of literature reporting consistent findings, semiquantitative evaluation of HRCT is associated with a number of well-documented limitations: it is 1) liable to significant interobserver variability; 2) poorly reproducible; 3) insensitive to subtle changes in disease extent over short follow-up periods; 4) time-consuming; and 5) requires domain expertise, which may not be available [7, 11, 32, 33]. This provides the rationale for applying computer-based image analysis to HRCT for both diagnostic support as well as reliable disease quantification, also known as quantitative computed tomography (QCT) (table 1).

QCT

Early studies

The earliest move toward QCT in pulmonary fibrosis used simple measures of lung density based on density masks or whole-lung HRCT histogram analysis [11]. Since the computed tomography (CT) histogram provides a graphical representation of lung density per voxel in a CT image, it allows the mean

TABLE 1 The tools of quantitative computed tomography (QCT) and deep learning

QCT	The computer is trained to identify and quantify patterns in HRCT. Its development requires "function engineering", a human operator
CALIPER	Uses volumetric local histogram and morphological analysis to characterise and quantify different HRCT patterns. Including the novel VRS variable, which has been shown to be an independent predictor of mortality and a potential tool to identify novel outcome-based radiologic phenotypes in various lung diseases
Adaptive multiple features method	Identifies and quantifies HRCT patterns based on textural analysis (normal lung, GGO, emphysema, honeycombing and nodules)
Quantitative lung fibrosis	Quantifies fibrotic reticular patterns. A total ILD extent composite of quantitative ILD is the sum of quantitative lung fibrosis, honeycombing and GGO patterns. It can provide complementary measures of disease progression to conventional lung physiology
Functional respiratory imaging	Combines low-dose HRCT with computer-based flow simulations. Functional respiratory imaging enables precise quantification of lung structure and function, with low variability for airway volumes, blood vessel volumes and airway resistances. In addition, it can evaluate airway volume, making it useful for measuring the severity of traction bronchiectasis, which is a predictor of mortality
Deep learning	Has the ability to autonomously identify patterns in high-dimensional data features (for example, HRCT scans). It has no human operator
SOFIA	The algorithm is trained to identify UIP-like features on HRCT. It provides a "UIP probability" score. It can predict disease progression and mortality in patients with suspected IPF
Data-driven texture analysis	Classifies image patches based on the presence of fibrosis and quantifies fibrosis extent on HRCT. It can stratify patients based on fibrosis extent

CALIPER: Computer-Aided Lung Informatics for Pathology Evaluation and Rating; SOFIA: Systematic Objective Fibrotic Imaging Analysis Algorithm; HRCT: high-resolution computed tomography; VRS: vascular-related structures; GGO: ground-glass opacities; ILD: interstitial lung disease; UIP: usual interstitial pneumonia; IPF: idiopathic pulmonary fibrosis.

lung attenuation, skewness and kurtosis to be calculated. Kurtosis describes the sharpness of the peak of the histogram, whereas skewness is a measure of the lack of symmetry of the CT histogram. Lung fibrosis increases the mean lung attenuation and reduces the kurtosis and leftward skewness of the histogram; therefore, these metrics may be used as surrogates of fibrosis extent on CT. In 144 IPF patients, Best *et al*. [34] reported a correlation between kurtosis and physiological decline and mortality. A key difficulty with this approach is that it cannot discriminate between different HRCT patterns commonly seen in patients with IPF. Ash *et al*. [35] described local histogram-based objective quantification of different radiologic patterns of disease in 46 patients with IPF and found strong correlations between visual and objective histogram-based scores for disease extent as well as a poor prognosis in patients with higher fibrosis and honeycombing extent scores.

Computer-Aided Lung Informatics for Pathology Evaluation and Rating

Computer-Aided Lung Informatics for Pathology Evaluation and Rating (CALIPER) has been used to predict survival and future physiological decline in patients with IPF, using a computer vision based technique based on volumetric local histogram and morphological analysis to characterise and quantify different HRCT patterns [11]. Furthermore, CALIPER extracts the pulmonary vessels and provides an estimation of the vessel volume, reported as a novel "vascular-related structures" (VRS) variable. In a landmark study in 283 patients with IPF, Jacob *et al.* [36] demonstrated on multivariable survival analysis, which included CALIPER and semiquantitative HRCT pattern scores, that only the computer-based variables independently predicted mortality, with VRS being the strongest predictor among them. In a subsequent study, published in 2018 [37], the same group used a VRS threshold for cohort enrichment in an IPF drug trial setting to reduce the IPF drug trial population size by 25%. Importantly, the VRS score identified a subset of patients in whom antifibrotic therapy reduced forced vital capacity (FVC) decline. It is important to understand from these data that CALIPER was not originally designed to evaluate the pulmonary vessels; this variable was generated as a by-product of the software image pre-processing pipeline, which extracts the lung parenchyma from the airways and vessels. This finding is early evidence that computer-based image analysis provides an opportunity to identify novel HRCT biomarkers, including those that may not be accessible visually. CALIPER has also been applied to CTD-FLD and FHP. In a cohort of 203 all-comers CTD-FLD, Jacob *et al.* [38] demonstrated that VRS was an independent predictor of mortality across all CTD-FLD subgroups. In addition, the authors stratified patients into three prognostically distinct groups based on CALIPER-related HRCT variables, demonstrating the potential of this technology to identify novel outcome-based radiologic phenotypes in CTD. Likewise, in a cohort of 135 patients with a diagnosis of FHP, the same group [39] demonstrated stronger associations between restrictive functional indices and CALIPER-defined total ILD extent than semiquantitative scores. In a subsequent study, the authors applied a VRS threshold to identify a subgroup of patients with IPF-like disease behaviour among 103 patients with FHP. Similar results have been reported applying CALIPER to patients with unclassifiable fibrotic lung disease [40].

Adaptive multiple features method

The adaptive multiple features method (AMFM) identifies and quantifies HRCT patterns based on textural analysis, including normal lung, ground-glass opacification (GGO), emphysema, honeycombing and nodules [11]. Initially, this method was used to differentiate normal lung from the lung with emphysema. In the late 1990s, Uppaluri *et al*. [41] compared AMFM with mean lung density (MLD) and histogram-based analysis and demonstrated high precision for the AMFM method in discriminating between normal and emphysematous lung. Later studies extended these experiments to patients with IPF and sarcoidosis, comparing the AMFM with these two methods to objectively characterise four groups of subjects: normal lung, emphysema, IPF and sarcoidosis. In all four groups, the AMFM method demonstrated superiority over MLD and histogram-based analysis [42]. In 2017, Salisbury *et al*. [43] demonstrated in 199 IPF patients enrolled in the Prednisone, Azathioprine, and N-Acetylcysteine: a Study that Evaluates Response in IPF (PANTHER-IPF) treatment trial, that baseline fibrosis (measured as ground-glass reticular opacities (GGR)) measured by AMFM predicts disease progression. Interestingly, changes in GGR only weakly correlated with FVC changes, suggesting that a combination of FVC change and GGR change, as measured by the AMFM software, may provide improved prognostic signal over either variable in isolation (figure 1).

Quantitative lung fibrosis

Quantitative lung fibrosis (QLF) quantifies fibrotic reticular patterns [11]. A total ILD extent composite of quantitative interstitial lung disease (QILD) is the sum of QLF, honeycombing and GGO patterns. QLF has been shown to correlate well with lung function measurement in ILD patients and has been used to evaluate disease progression in IPF and scleroderma-related ILD treatment trials [44]. In a study of cyclophosphamide *versus* mycophenolate in 142 patients with scleroderma related ILD, Goldin *et al*. [45]

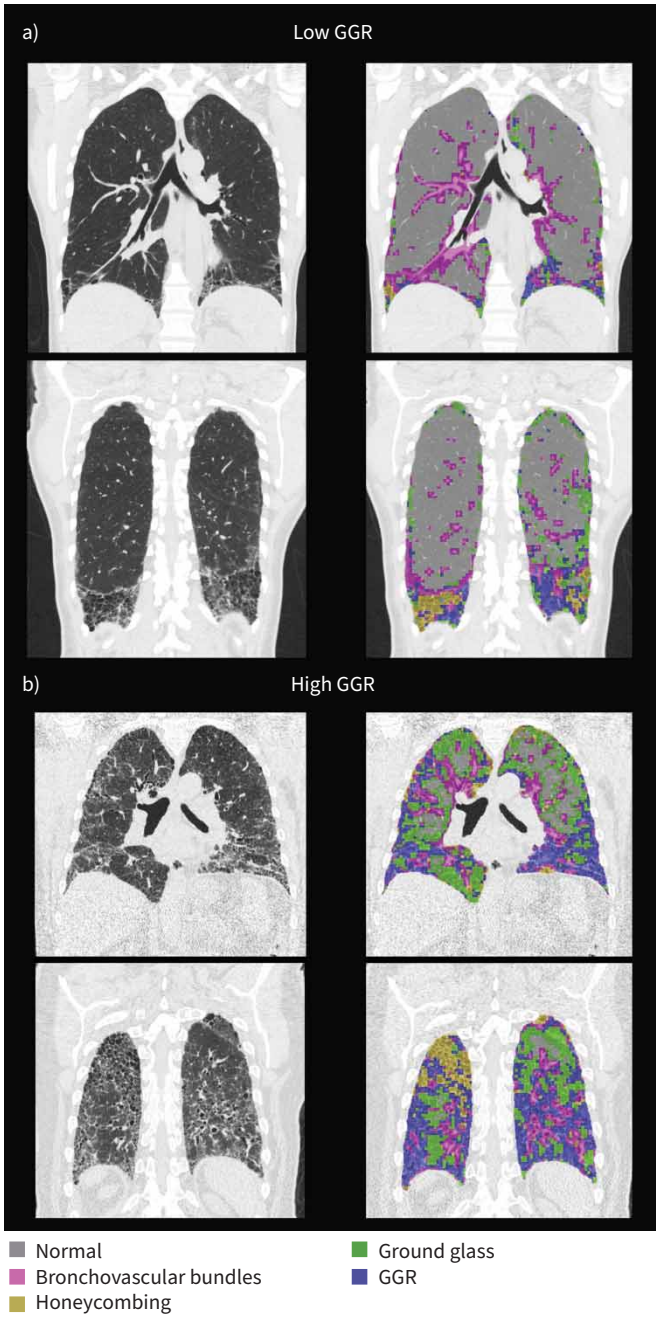

FIGURE 1 Adaptive multiple features method. a) A patient with low ground-glass reticular (GGR) texture and b) a patient with high GGR. Courtesy of Eric Hoffman (University of Iowa Caver College of Medicine, Iowa City, IA, USA).

found that QLF scores did not change in the treatment arms of the study, while QILD scores did show a small improvement in both treatment arms. The incorporation of QLF/QILD scores in secondary outcomes of clinical trials demonstrates the utility of computer-based imaging analysis tools for providing complementary measures of disease progression to conventional lung physiology (*i.e.* FVC) [46, 47] (figure 2).

Functional respiratory imaging

Functional respiratory imaging (FRI) combines low-dose HRCT with computer-based flow simulations. Respiratory gating using a handheld spirometer is performed during the acquisition to ensure repeatable lung volumes (figure 3). FRI allows regional quantification of lung structure and function and shows low

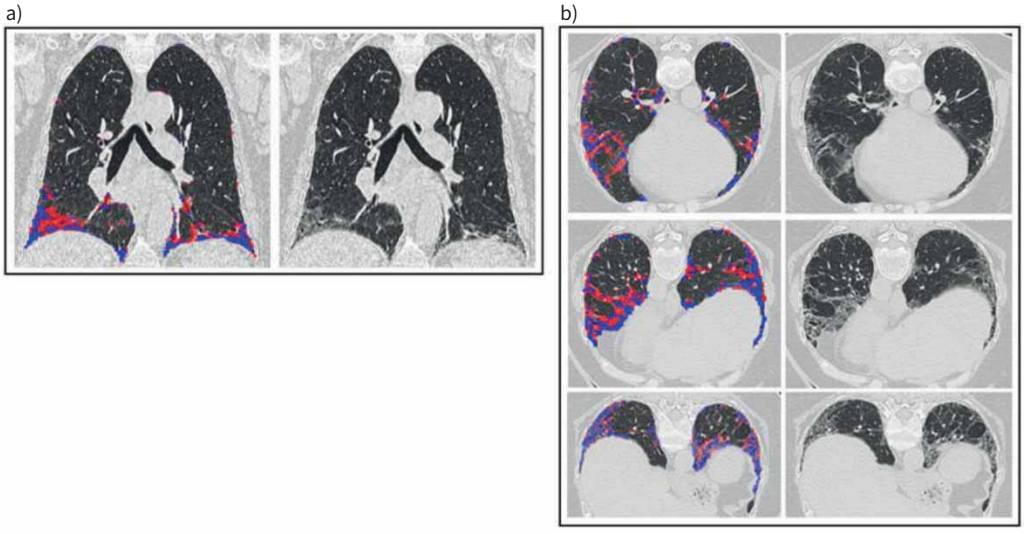

FIGURE 2 Coronal and axial computed tomography (CT) images with quantitative lung fibrosis (QLF) characterisation. a) Coronal CT with the classification of QLF (left) and original coronal CT image (right). b) Annotated axial high-resolution CT images with the classification of QLF (blue and red) and the corresponding original images. In whole lung, QLF extent is 10.6% and QLF score is 393 mL in volume. QLF scores in the right and left lungs are 11.5% and 9.5%, respectively. QLF scores were 20.1% and 19.7% in the right and left lower lobes (142 mL and 105 mL), respectively. The QLF score quantifies the extent and characterises the distribution of pulmonary fibrosis as predominantly lower lung disease. Courtesy of Grace Hyun J. Kim and Jonathan G. Goldin (University of California (UCLA), Los Angeles, CA, USA).

variability (1–3%) for airway volumes, blood vessel volumes and airway resistances [48]. In addition, FRI can assess airway volume and therefore can quantify the severity of traction bronchiectasis, a potent predictor of mortality based on several studies that applied semiquantitative airway assessments. Studies in

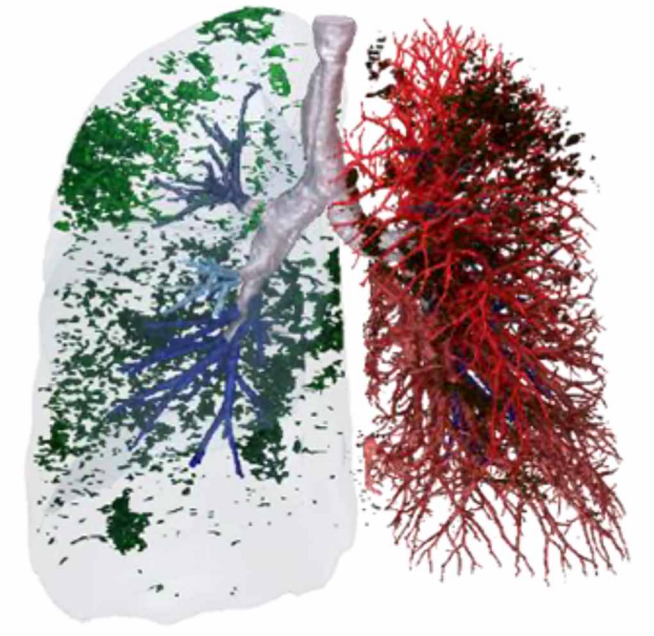

FIGURE 3 Functional respiratory imaging. Visualisation and quantification of airway volumes (blue), lobe volume, fibrosis (green), emphysema (black) and blood vessel volume (red). Reproduced with permission from Fluidda (Kontich, Belgium).

IPF show that disease progression, as determined by FVC decline, is associated with a reduction in CT-measured lung volumes (R^2=0.80, p<0.001) and an increase in relative airway volumes (R^2=0.29, p<0.001). Changes in FVC are correlated with changes in lung volumes (R^2=0.18, p<0.001) and changes in relative airway calibre (R^2=0.15, p<0.001) [49]. Lobe and airway volumes can be significantly affected by IPF, whereas conventional measures such as FVC remain within the normal (healthy) range, while FRI metrics capture early changes. Additional studies are needed to determine minimal clinically important differences.

Deep learning

A key drawback of many of the QCT tools described is that their development requires some degree of "feature engineering": the computer is trained to identify and quantify specific HRCT patterns by human operators. This means that, in principle, all the limitations associated with visual HRCT assessment are incorporated into the system. A second significant issue is that the features upon which the computer is trained need to be known *a priori*, negating the possibility that novel, visually inaccessible HRCT biomarkers might be discovered. Both of these challenges can be overcome if the computer can learn to extract the most predictive features from the images in an autonomous fashion. This is the key advantage of deep learning.

Deep learning is a form of machine learning that has the capacity to autonomously identify patterns in high-dimensional data (*e.g.* HRCT scans) and map these patterns to end-points such as diagnosis and future disease progression [7, 50–52]. Deep learning is very efficient at identifying subtle features within images that are important while at the same time ignoring irrelevant variations between images including those introduced by different HRCT techniques. The key advantage of deep learning over many existing QCT techniques is that it simultaneously optimises feature extraction and classification during algorithm training; *a priori* knowledge of what image features to quantify for a given classification problem is not necessary. More concretely, deep learning bypasses the need to train computers on specific patterns; the computer learns itself, during training, which patterns on HRCT are most important for predicting a given task. In addition, this approach has the added advantage of avoiding all the limitations associated with visual HRCT assessment. Perhaps most importantly, since the computer learns autonomously without explicit programming, an opportunity is created for the identification of novel HRCT biomarkers, including those that are not readily identified visually. In respiratory medicine, deep learning has been applied successfully to lung cancer detection, predicting mortality in patients with COPD and classifying fibrotic lung disease on CT scans [7, 50, 53].

Applications of deep learning to fibrotic lung disease

In principle, deep learning can be applied to a number of unresolved research questions related to imaging in fibrotic lung disease. Two important unanswered questions relate to 1) predicting progressive fibrotic lung disease using baseline imaging and clinical data; and 2) early detection of clinically significant fibrotic lung disease.

Identifying patients with progressive fibrotic lung disease

The reliable identification of progressive fibrotic lung disease using baseline imaging and clinical data is of immediate clinical importance [8–10, 54–58]. Since antifibrotic therapy is currently only licensed in those patients who demonstrate progression (*i.e.* progressive pulmonary fibrosis), patients must first undergo a period of progression before they qualify for treatment, meaning that an opportunity to initiate early treatment and reduce functional decline is missed. Based on published data from recent clinical trials, UIP and probable UIP (UIP-like disease) in general exhibit progressive disease behaviour, but the progressive disease is not confined to patients with UIP-like disease; currently, we cannot accurately predict progression using baseline HRCT data in this non-UIP group [8, 59].

Recently, a deep learning algorithm, Systematic Objective Fibrotic Imaging Analysis Algorithm (SOFIA), trained to identify UIP-like features on HRCT and provide a "UIP probability" score was used to predict progression in a cohort of 504 suspected IPF patients, drawn from the Australian IPF Registry [7]. This novel HRCT biomarker, the UIP probability score, was predictive of mortality, independently of disease severity (when expressed as a total fibrosis score on HRCT, or lung function). Furthermore, on subgroup analysis in patients whose HRCT was considered indeterminate (*i.e.* the HRCT was considered unhelpful based on visual assessment by two expert thoracic radiologists), the UIP probability score was again a strong predictor of mortality (hazard ratio (HR) 1.73, 95% CI 1.40–2.14; p<0.0001). Finally, in patients who underwent surgical lung biopsy (n=86), the UIP probability score predicted mortality independently of guideline-based histologic diagnosis and total fibrosis extent, with both these latter variables failing to reach statistical significance (HR 1.75, 95% CI 1.37–2.25; p<0.0001). It is important to point out that

radiologists can also provide a UIP probability score, and this outperforms guideline-based HRCT diagnosis in survival analysis [7]. However, in this setting, radiologists tend to default to the extremes of this scale (*i.e.* they tend to assign a UIP probability of 0% or 100%), whereas SOFIA provides a granular probability score as a continuous variable, regardless of the HRCT pattern; subjective biases to which human assessment are vulnerable do not exist (figure 4).

It is important to highlight that further work is needed to decode the outputs of SOFIA, particularly in cases where there is significant disagreement between the algorithm and radiologists. More generally, a key challenge in deep learning is that the complexity that makes neural networks so efficient at identifying patterns in large datasets can also make them difficult to interpret. Neural networks are often regarded as "black boxes", which is viewed as an obstacle to their implementation. Explainability is an increasingly important component of algorithm development, particularly when algorithmic decision-making is based on features contained within the images that are invisible to human observers. In addition, efficient deep learning relies on being able to understand why an algorithm misclassifies certain images, making algorithm interpretability crucial.

Deep learning based QCT has been developed. Data-driven texture analysis (DTA) is a deep learning based tool which utilises a convolutional neural network to classify image patches based on the presence of fibrosis and quantifies fibrosis extent on HRCT. DTA fibrosis score has demonstrated good correlation with lung function and visual quantification of fibrosis by experts and can stratify patients based on fibrosis extent (figure 5). By quantifying baseline line fibrosis extent, it can also be used to predict disease progression (HR 1.14, 95% CI 1.08–1.19; p<0.0001) [60–62]. Humphries *et al.* [62] reported significant associations with FVC and diffusing capacity of the lung for carbon monoxide decline in a cohort of 393 IPF patients, as well as statistically significant outcome prediction, independent of lung function.

Detection of early fibrotic lung disease

The second open research question to which deep learning can be applied is the characterisation of interstitial lung abnormalities (ILAs). ILAs are defined as interstitial abnormalities that exceed 5% extent of the total lung volume on HRCT, and they present thorny clinical problem. Data extracted from longitudinal lung cancer and cardiovascular cohort studies show shared clinical and genetic associations between incidentally detected ILAs on HRCT and IPF. ILAs are associated with ageing and are more commonly seen in smokers. ILAs are also seen in those expressing MUC5B promoter polymorphism positivity [63, 64] and ILA progression correlates to physiological decline. However, ILAs are common, seen in 7–9% of lung cancer screening subjects, exceeding the prevalence of IPF by almost two orders of magnitude [65]. The current challenge is that it is not possible to predict which ILAs will progress to

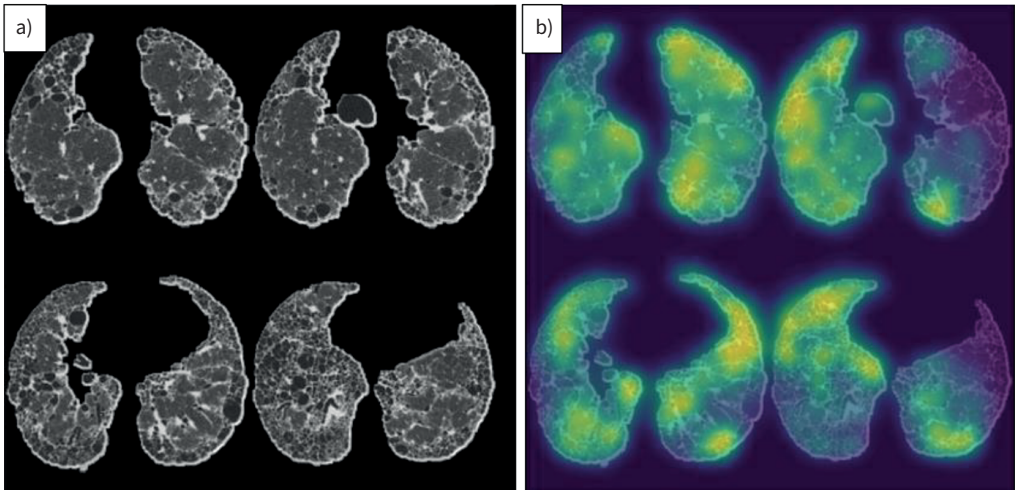

FIGURE 4 Systematic Objective Fibrotic Lung Disease Analysis Algorithm (SOFIA). a) Four-slice high-resolution computed tomography montage of segmented lung slices depicted peripheral honeycombing consistent with a usual interstitial pneumonia (UIP) pattern. SOFIA scores for this case were UIP 0.9963; probable UIP 0.0036; indeterminate 0.0001; and alternative diagnosis 0.000. b) Saliency map for a), highlighting regions within the montage that were the most influential in algorithmic decision-making.

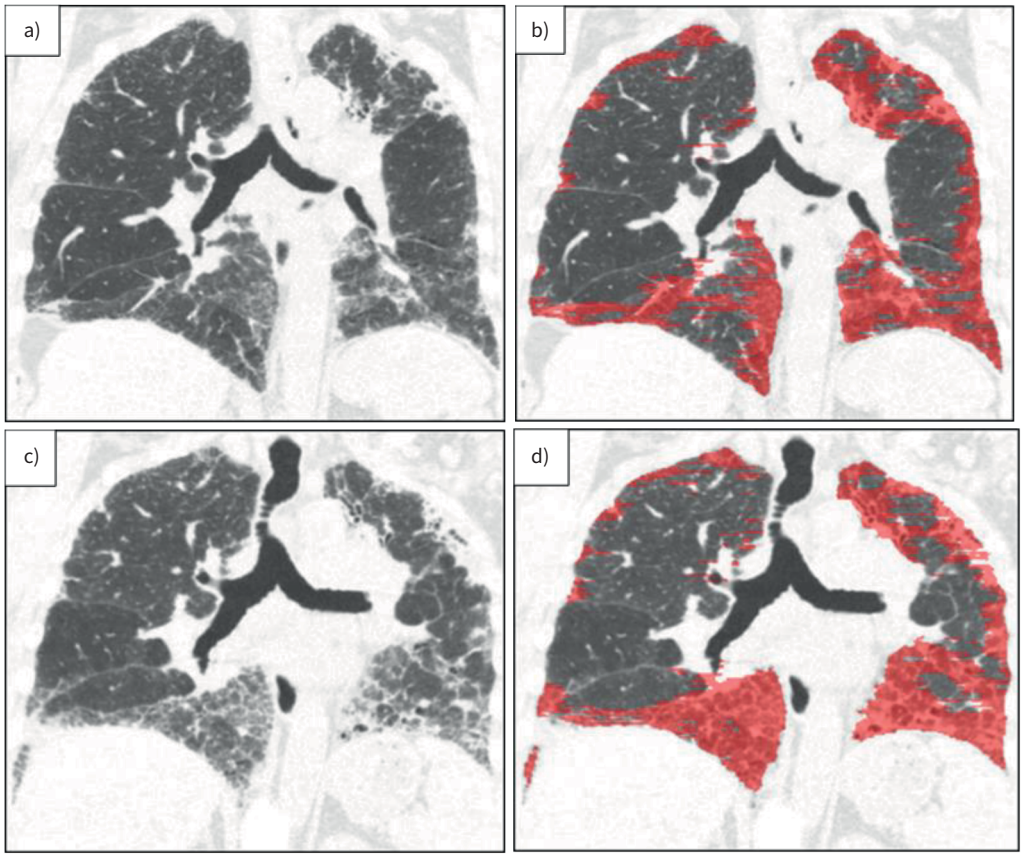

FIGURE 5 Data-driven texture analysis (DTA). Coronal computed tomography (CT) sections on a 66-year-old female with idiopathic pulmonary fibrosis. a) Visual CT pattern was indeterminate for usual interstitial pneumonia. Baseline CT with b) DTA classification as red overlay. DTA score (calculated as percentage of lung volume classified as fibrosis) was 33.0 at baseline. c) Follow-up CT at 1 year and d) DTA classification as red overlay. DTA score increased to 39.0 at 1 year follow-up. Courtesy of Stephen M. Humphries (National Jewish Health, Denver, CO, USA).

clinically significant fibrotic lung disease and which will not. As with diagnosis in established fibrotic lung disease, the current ILA classification is based on visually defined morphology, rather than disease behaviour, which means that classification of incidentally identified ILAs is associated with all the limitations associated with visual HRCT evaluation. Furthermore, the current ILA definition represents an umbrella term encompassing a range of nonfibrotic and fibrotic patterns. This definition will need refinement if progressive ILAs are to be identified reliably. As with predicting progressive behaviour when fibrosis is established, one solution might be found in deep learning based analysis; algorithmic training could be anchored to ILA behaviour with no *a priori* assumptions as to the importance of individual ILA patterns. A major challenge to this approach will be the collating of sufficiently large datasets to adequately power algorithm training.

Challenges to development and implementation

The use of QCT as a biomarker in fibrotic lung disease faces several barriers. These include access to high-quality data in sufficient quantities to drive novel QCT development; recognising and minimising biases in algorithm training; improving algorithm explainability; ensuring equal access for patients to artificial intelligence (AI)-based technology; and establishing reference standards for training, testing and algorithm deployment.

The availability of large and diverse datasets is a critical factor in the development of effective machine learning models. Open-source datasets such as the Open Source Imaging Consortium (OSIC; www.osicild.org) can help address these limitations by making data more accessible and secure, while also addressing privacy and ethical concerns. The multidisciplinary nature of OSIC, engaging radiologists, clinicians and computer

and data scientists, as well as industry stakeholders helps to ensure the credibility and trustworthiness of the dataset, thus making it a valuable resource for the development of AI-powered healthcare solutions.

The integration of machine learning with pathogenetics can have a major impact on drug development. Machine learning can help identify patterns and correlations in large population data, allowing the testing of hypotheses on a larger scale. This can lead to more personalised and effective treatments, as well as a deeper understanding of disease mechanisms. By leveraging the power of machine learning, drug development can be more efficient and targeted, ultimately improving patient outcomes.

Deep learning algorithms come with unique risks, because of they can reinforce biases in training data. Missing or unbalanced data can affect algorithm performance and amplify inequalities in healthcare in ways that are difficult to detect. Subgroups of patients with rare diseases may not see the benefit of these AI-based imaging analysis techniques because of insufficient data for algorithm development [66]. In addition, deep learning algorithms may be manipulated to output conclusions that trend toward the use of specific third-party tests. Establishing ethical frameworks with buy-in from all stakeholders (and in particular, patients) will be needed to foster trust in this technology. Bespoke governance frameworks that are tailored to address the unique challenges associated with AI will probably be needed. Preserving trust and transparency will be of paramount importance. Finding ways to encode ethical standards into AI training will be essential, as well as preserving trust and transparency.

Encouraging the medical community to fully embrace AI and machine learning tools may be hampered by a lack of understanding and concerns about quality, safety and accuracy. However, it is important to consider that first the quantitative analysis provided by these tools can offer more reliable and objective data for disease assessment and precision medicine [67–71]. Second, this can aid in clinical decision-making and improve the accuracy of predictions about disease progression. It will also be important for all stakeholders to receive appropriate education and training on the use of these tools and how to appraise and overcome their limitations.

Conclusion

Increasingly, QCT and AI are recognised as valuable tools in the diagnosis and prognosis of ILDs. Two key advantages are 1) they offer the advantage of being more precise and efficient compared to semiquantitative methods; and 2) they can help in decision-making for physicians. However, there are still challenges in terms of acceptance by the medical community and the navigation of technical and bureaucratic hurdles.

Provenance: Commissioned article, peer reviewed.

Conflict of interest: F.N. Felder reports payment or honoraria for lectures, presentations, speakers' bureaus, manuscript writing or educational events from Boehringer Ingelheim, outside the submitted work. S.L.F. Walsh reports payment or honoraria for lectures, presentations, speakers' bureaus, manuscript writing or educational events from Boehringer Ingelheim, Roche and Galapagos, outside the submitted work.

References

1 Wijsenbeek M, Suzuki A, Maher TM. Interstitial lung diseases. *Lancet* 2022; 400: 769–786.
2 Lederer DJ, Martinez FJ. Idiopathic pulmonary fibrosis. *N Engl J Med* 2018; 379: 797–798.
3 Raghu G, Collard HR, Egan JJ, *et al.* An official ATS/ERS/JRS/ALAT statement: idiopathic pulmonary fibrosis: evidence-based guidelines for diagnosis and management. *Am J Respir Crit Care Med* 2011; 183: 788–824.
4 Grunewald J, Grutters JC, Arkema EV, *et al.* Sarcoidosis. *Nat Rev Dis Primers* 2019; 5: 45.
5 Spagnolo P, Rossi G, Trisolini R, *et al.* Pulmonary sarcoidosis. *Lancet Respir Med* 2018; 6: 389–402.
6 Costabel U, Miyazaki Y, Pardo A, *et al.* Hypersensitivity pneumonitis. *Nat Rev Dis Primers* 2020; 6: 65.
7 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.
8 Flaherty KR, Wells AU, Cottin V, *et al*. Nintedanib in progressive fibrosing interstitial lung diseases. *N Engl J Med* 2019;381: 1718–1727.
9 Takei R, Brown KK, Yamano Y, *et al*. Prevalence and prognosis of chronic fibrosing interstitial lung diseases with a progressive phenotype. *Respirology* 2022; 27: 333–340.
10 Maher TM, Corte TJ, Fischer A, *et al*. Pirfenidone in patients with unclassifiable progressive fibrosing interstitial lung disease: a double-blind, randomised, placebo-controlled, phase 2 trial. *Lancet Respir Med* 2020; 8: 147–157.

11 Wu X, Kim GH, Salisbury ML, *et al.* Computed tomographic biomarkers in idiopathic pulmonary fibrosis. The future of quantitative analysis. *Am J Respir Crit Care Med* 2019; 199: 12–21.
12 Richeldi L, du Bois RM, Raghu G, *et al.* Efficacy and safety of nintedanib in idiopathic pulmonary fibrosis. *N Engl J Med* 2014; 370: 2071–2082.
13 Flaherty KR, Mumford JA, Murray S, *et al.* Prognostic implications of physiologic and radiographic changes in idiopathic interstitial pneumonia. *Am J Respir Crit Care Med* 2003; 168: 543–548.
14 Kim EJ, Elicker BM, Maldonado F, *et al.* Usual interstitial pneumonia in rheumatoid arthritis-associated interstitial lung disease. *Eur Respir J* 2010; 35: 1322–1328.
15 Salisbury ML, Gu T, Murray S, *et al.* Hypersensitivity pneumonitis: radiologic phenotypes are associated with distinct survival time and pulmonary function trajectory. *Chest* 2019; 155: 699–711.
16 Choe J, Chae EJ, Kim YJ, *et al.* Serial changes of CT findings in patients with chronic hypersensitivity pneumonitis: imaging trajectories and predictors of fibrotic progression and acute exacerbation. *Eur Radiol* 2021; 31: 3993–4003.
17 Raghu G, Remy-Jardin M, Myers JL, *et al.* Diagnosis of idiopathic pulmonary fibrosis. An official ATS/ERS/JRS/ALAT clinical practice guideline. *Am J Respir Crit Care Med* 2018; 198: e44–e68.
18 Gay SE, Kazerooni EA, Toews GB, *et al.* Idiopathic pulmonary fibrosis: predicting response to therapy and survival. *Am J Respir Crit Care Med* 1998; 157: 1063–1072.
19 Walsh SL, Sverzellati N, Devaraj A, *et al.* Chronic hypersensitivity pneumonitis: high resolution computed tomography patterns and pulmonary function indices as prognostic determinants. *Eur Radiol* 2012; 22: 1672–1679.
20 Walsh SL, Sverzellati N, Devaraj A, *et al.* Connective tissue disease related fibrotic lung disease: high resolution computed tomographic and pulmonary function indices as prognostic determinants. *Thorax* 2014; 69: 216–222.
21 Mogulkoc N, Brutsche MH, Bishop PW, *et al.* Pulmonary function in idiopathic pulmonary fibrosis and referral for lung transplantation. *Am J Respir Crit Care Med* 2001; 164: 103–108.
22 Sumikawa H, Johkoh T, Colby TV, *et al.* Computed tomography findings in pathological usual interstitial pneumonia: relationship to survival. *Am J Respir Crit Care Med* 2008; 177: 433–439.
23 Lynch DA, Godwin JD, Safrin S, *et al.* High-resolution computed tomography in idiopathic pulmonary fibrosis: diagnosis and prognosis. *Am J Respir Crit Care Med* 2005; 172: 488–493.
24 Edey AJ, Devaraj AA, Barker RP, *et al.* Fibrotic idiopathic interstitial pneumonias: HRCT findings that predict mortality. *Eur Radiol* 2011; 21: 1586–1593.
25 Jacob J, Aksman L, Mogulkoc N, *et al.* Serial CT analysis in idiopathic pulmonary fibrosis: comparison of visual features that determine patient outcome. *Thorax* 2020; 75: 648–654.
26 Robbie H, Wells AU, Fang C, *et al.* Serial decline in lung volume parameters on computed tomography (CT) predicts outcome in idiopathic pulmonary fibrosis (IPF). *Eur Radiol* 2022; 32: 2650–2660.
27 Chung JH, Zhan X, Cao M, *et al.* Presence of air trapping and mosaic attenuation on chest computed tomography predicts survival in chronic hypersensitivity pneumonitis. *Ann Am Thorac Soc* 2017; 14: 1533–1538.
28 Goh NS, Desai SR, Veeraraghavan S, *et al.* Interstitial lung disease in systemic sclerosis: a simple staging system. *Am J Respir Crit Care Med* 2008; 177: 1248–1254.
29 Ley B, Elicker BM, Hartman TE, *et al.* Idiopathic pulmonary fibrosis: CT and risk of death. *Radiology* 2014; 273: 570–579.
30 Walsh SL, Wells AU, Sverzellati N, *et al.* An integrated clinicoradiological staging system for pulmonary sarcoidosis: a case–cohort study. *Lancet Respir Med* 2014; 2: 123–130.
31 Wells AU, Antoniou KM. The prognostic value of the GAP model in chronic interstitial lung disease: the quest for a staging system. *Chest* 2014; 145: 672–674.
32 Walsh SL, Calandriello L, Sverzellati N, *et al.* Interobserver agreement for the ATS/ERS/JRS/ALAT criteria for a UIP pattern on CT. *Thorax* 2016; 71: 45–51.
33 Widell J, Lidén M. Interobserver variability in high-resolution CT of the lungs. *Eur J Radiol Open* 2020; 7: 100228.
34 Best AC, Meng J, Lynch AM, *et al.* Idiopathic pulmonary fibrosis: physiologic tests, quantitative CT indexes, and CT visual scores as predictors of mortality. *Radiology* 2008; 246: 935–940.
35 Ash SY, Harmouche R, Vallejo DL, *et al.* Densitometric and local histogram based analysis of computed tomography images in patients with idiopathic pulmonary fibrosis. *Respir Res* 2017; 18: 45.
36 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.
37 Jacob J, Bartholmai BJ, Rajagopalan S, *et al.* Predicting outcomes in idiopathic pulmonary fibrosis using automated computed tomographic analysis. *Am J Respir Crit Care Med* 2018; 198: 767–776.
38 Jacob J, Bartholmai BJ, Rajagopalan S, *et al.* Evaluation of computer-based computer tomography stratification against outcome models in connective tissue disease-related interstitial lung disease: a patient outcome study. *BMC Med* 2016; 14: 190.

39 Jacob J, Bartholmai BJ, Brun AL, *et al.* Evaluation of visual and computer-based CT analysis for the identification of functional patterns of obstruction and restriction in hypersensitivity pneumonitis. *Respirology* 2017; 22: 1585–1591.
40 Jacob J, Bartholmai BJ, Rajagopalan S, *et al.* Unclassifiable-interstitial lung disease: outcome prediction using CT and functional indices. *Respir Med* 2017; 130: 43–51.
41 Uppaluri R, Mitsa T, Sonka M, *et al.* Quantification of pulmonary emphysema from lung computed tomography images. *Am J Respir Crit Care Med* 1997; 156: 248–254.
42 Uppaluri R, Hoffman EA, Sonka M, *et al.* Interstitial lung disease: a quantitative study using the adaptive multiple feature method. *Am J Respir Crit Care Med* 1999; 159: 519–525.
43 Salisbury ML, Lynch DA, van Beek EJ, *et al.* Idiopathic pulmonary fibrosis: the association between the adaptive multiple features method and fibrosis outcomes. *Am J Respir Crit Care Med* 2017; 195: 921–929.
44 Kim HG, Tashkin DP, Clements PJ, *et al.* A computer-aided diagnosis system for quantitative scoring of extent of lung fibrosis in scleroderma patients. *Clin Exp Rheumatol* 2010; 28: Suppl. 62, S26–S35.
45 Goldin JG, Kim GHJ, Tseng CH, *et al.* Longitudinal changes in quantitative interstitial lung disease on computed tomography after immunosuppression in the Scleroderma Lung Study II. *Ann Am Thorac Soc* 2018; 15: 1286–1295.
46 Richeldi L, Fernández Pérez ER, Costabel U, *et al.* Pamrevlumab, an anti-connective tissue growth factor therapy, for idiopathic pulmonary fibrosis (PRAISE): a phase 2, randomised, double-blind, placebo-controlled trial. *Lancet Respir Med* 2020; 8: 25–33.
47 Khanna D, Lin CJF, Furst DE, *et al.* Tocilizumab in systemic sclerosis: a randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet Respir Med* 2020; 8: 963–974.
48 Burrowes KS, De Backer J, Kumar H. Image-based computational fluid dynamics in the lung: virtual reality or new clinical practice? *Wiley Interdiscip Rev Syst Biol Med* 2017; 9: e1392.
49 Clukers J, Lanclus M, Mignot B, *et al.* Quantitative CT analysis using functional imaging is superior in describing disease progression in idiopathic pulmonary fibrosis compared to forced vital capacity. *Respir Res* 2018; 19: 213.
50 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.
51 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.
52 Yu W, Zhou H, Choi Y, *et al.* Multi-scale, domain knowledge-guided attention + random forest: a two-stage deep learning-based multi-scale guided attention models to diagnose idiopathic pulmonary fibrosis from computed tomography images. *Med Phys* 2023; 50: 894–905.
53 Wells AU, Walsh SLF. Quantitative computed tomography and machine learning: recent data in fibrotic interstitial lung disease and potential role in pulmonary sarcoidosis. *Curr Opin Pulm Med* 2022; 28: 492–497.
54 Gagliardi M, Berg DV, Heylen CE, *et al.* Real-life prevalence of progressive fibrosing interstitial lung diseases. *Sci Rep* 2021; 11: 23988.
55 Simpson T, Barratt SL, Beirne P, *et al.* The burden of progressive fibrotic interstitial lung disease across the UK. *Eur Respir J* 2021; 58: 2100221.
56 Nasser M, Larrieu S, Si-Mohamed S, *et al.* Progressive fibrosing interstitial lung disease: a clinical cohort (the PROGRESS study). *Eur Respir J* 2021; 57: 2002718.
57 Khor YH, Farooqi M, Hambly N, *et al.* Patient characteristics and survival for progressive pulmonary fibrosis using different definitions. *Am J Respir Crit Care Med* 2023; 207: 102–105.
58 Pugashetti JV, Adegunsoye A, Wu Z, *et al.* Validation of proposed criteria for progressive pulmonary fibrosis. *Am J Respir Crit Care Med* 2023; 207: 69–76.
59 Behr J, Prasse A, Kreuter M, *et al.* Pirfenidone in patients with progressive fibrotic interstitial lung diseases other than idiopathic pulmonary fibrosis (RELIEF): a double-blind, randomised, placebo-controlled, phase 2b trial. *Lancet Respir Med* 2021; 9: 476–486.
60 Humphries SM, Yagihashi K, Huckleberry J, *et al.* Idiopathic pulmonary fibrosis: data-driven textural analysis of extent of fibrosis at baseline and 15-month follow-up. *Radiology* 2017; 285: 270–278.
61 Humphries SM, Swigris JJ, Brown KK, *et al.* Quantitative high-resolution computed tomography fibrosis score: performance characteristics in idiopathic pulmonary fibrosis. *Eur Respir J* 2018; 52: 1801384.
62 Humphries SM, Mackintosh JA, Jo HE, *et al.* Quantitative computed tomography predicts outcomes in idiopathic pulmonary fibrosis. *Respirology* 2022; 27: 1045–1053.
63 Kliment CR, Araki T, Doyle TJ, *et al.* A comparison of visual and quantitative methods to identify interstitial lung abnormalities. *BMC Pulm Med* 2015; 15: 134.
64 Hatabu H, Hunninghake GM, Richeldi L, *et al.* Interstitial lung abnormalities detected incidentally on CT: a position paper from the Fleischner Society. *Lancet Respir Med* 2020; 8: 726–737.
65 Jin GY, Lynch D, Chawla A, *et al.* Interstitial lung abnormalities in a CT lung cancer screening population: prevalence and progression rate. *Radiology* 2013; 268: 563–571.

66 Willemink MJ, Koszek WA, Hardell C, *et al.* Preparing medical imaging data for machine learning. *Radiology* 2020; 295: 4–15.
67 Strohm L, Hehakaya C, Ranschaert ER, *et al.* Implementation of artificial intelligence (AI) applications in radiology: hindering and facilitating factors. *Eur Radiol* 2020; 30: 5525–5532.
68 Huisman M, Ranschaert E, Parker W, *et al.* An international survey on AI in radiology in 1041 radiologists and radiology residents part 2: expectations, hurdles to implementation, and education. *Eur Radiol* 2021; 31: 8797–8806.
69 Reyes M, Meier R, Pereira S, *et al.* On the interpretability of artificial intelligence in radiology: challenges and opportunities. *Radiol Artif Intell* 2020; 2: e190043.
70 Cutillo CM, Sharma KR, Foschini L, *et al.* Machine intelligence in healthcare – perspectives on trustworthiness, explainability, usability, and transparency. *NPJ Digit Med* 2020; 3: 47.
71 Vellido A. Societal issues concerning the application of artificial intelligence in medicine. *Kidney Dis* 2019; 5: 11–17.