



# Radiological diagnosis of interstitial lung disease: is it all about pattern recognition?

Simon L.F. Walsh<sup>1</sup> and Martin Kolb<sup>2</sup>

**Affiliations:** <sup>1</sup>Dept of Radiology, King's College Hospital Foundation Trust, London, UK. <sup>2</sup>Departments of Medicine and Pathology/Molecular Medicine, McMaster University, Firestone Institute for Respiratory Health, St Joseph's Healthcare, Hamilton, ON, Canada.

**Correspondence:** Simon L.F. Walsh, Dept of Radiology, King's College Hospital Foundation Trust, Denmark Hill, Brixton, London SE5 9RS, UK. E-mail: slfwalsh@gmail.com

 @ERSpublications

**Radiological diagnosis of ILD is pattern-based and linked to underlying histology. The future of radiological diagnosis in ILD may be the identification of disease behaviour-based radiological phenotypes that predict disease outcome.** <http://ow.ly/Dg0Q30l4Rh8>

**Cite this article as:** Walsh SLF, Kolb M. Radiological diagnosis of interstitial lung disease: is it all about pattern recognition? *Eur Respir J* 2018; 52: 1801321 [<https://doi.org/10.1183/13993003.01321-2018>].

Hypersensitivity pneumonitis (HP) is a complex fibroinflammatory lung condition that arises from repeated exposure, usually to aerosolised organic antigens, in sensitised individuals. Although HP is a well-recognised clinical entity, the underlying mechanisms that drive disease progression are poorly understood.

Making a diagnosis of HP depends on the presence of variable combinations of clinical features, including the presence of serum antibodies to inciting antigens, lymphocytosis on bronchoalveolar lavage, compatible features on high-resolution computed tomography (HRCT) and if available, the presence of loosely formed granuloma in a bronchiolocentric location on lung biopsy [1]. However, in many cases, making a confident diagnosis of HP is hampered by marginal test results, nonspecific HRCT features and perhaps most importantly, the lack of internationally agreed diagnostic guidelines for the disease. These difficulties were brought into sharp relief by a recent study of multidisciplinary practice that reported miserable diagnostic agreement (weighted  $\kappa$ -coefficient ( $\kappa_w$ )=0.29) between expert multidisciplinary groups assigning a diagnosis of HP to a set of standardised cases drawn from a tertiary referral centre for diffuse lung diseases [2]. In contrast, interobserver agreement for idiopathic pulmonary fibrosis (IPF) was good ( $\kappa_w$ =0.71), reflecting the positive impact evidence-based guidelines have on diagnostic performance. Adding to these challenges, HP is highly variable in its presentation, response to antigen avoidance and treated course in individual patients. The response to the "HP challenge" has been a groundswell of research focused on disentangling the complexities of HP diagnosis, all of which aim squarely at developing a case definition for HP that can be readily applied in routine clinical practice. This effort has resulted in diagnostic algorithms for HP that combine salient clinical variables with HRCT features and clinical perspectives, recommending different but similar diagnostic approaches to HP [3–5].

HRCT plays a central role in the evaluation of patients with diffuse lung diseases and often has a significant impact on subsequent management decisions including the need for lung biopsy. Historically, radiological diagnosis has been inextricably linked to histopathology. The first systematic studies of computed tomography in diffuse lung disease began in the mid-1980s and the development of HRCT in the early 1990s was followed by more than 10 years of HRCT-pathology correlative studies that bridged

---

Received: July 15 2018 | Accepted: July 16 2018

Copyright ©ERS 2018

the gap between the microscopic world of interstitial lung disease (ILD) and their macroscopic appearances on HRCT. This era, typified by the meticulous work of the radiologist–pathologist duo of Nestor Müller and Roberta Miller, formed the foundations upon which HRCT diagnosis stands today. During that period, the primacy of histopathology as a diagnostic test in ILD was reflected in two ways: first, the official American Thoracic Society (ATS)/European Respiratory Society (ERS) classification of idiopathic interstitial pneumonias (IIPs) at the time was based on histopathological morphology; and second, histopathology was the diagnostic reference standard against which the veracity of all other diagnostic tests, including HRCT, was gauged.

As with other ILDs, the HRCT appearances of “biopsy-proven” HP were studied during this period. Early work highlighted the importance of poorly defined ground-glass nodules, heterogeneous or peribronchovascular axial distribution of disease, mosaic attenuation, and diffuse ground-glass opacification as important features of HP on HRCT [6–8]. Next, SILVA *et al.* [9] evaluated the accuracy of HRCT for distinguishing chronic fibrotic hypersensitivity pneumonitis (cHP) from IPF and nonspecific interstitial pneumonia (NSIP); this is the classic imaging conundrum faced by practising ILD radiologists in multidisciplinary meetings the world over. In this highly cited study, the authors concluded that lobular areas of air trapping, centrilobular nodules and an absence of lower zone predominance predicted a clinical diagnosis of cHP, but also highlighted that separating cHP from IPF and NSIP is possible in only ~50% of cases.

In a study presented in this issue of the *European Respiratory Journal*, SALISBURY *et al.* [10] add to the growing HP literature by developing, for the first time, a “rule-in” radiological diagnosis model for HP. First, they assembled a diverse cohort of ILD patients (n=356) who underwent diagnostic case review at their local multidisciplinary ILD conference. Additional HP patients were identified by their electronic medical records. Importantly, the outcome for analysis was clinical diagnosis following multidisciplinary discussion (or documented diagnosis in the clinical records), and HRCT in isolation was not used to verify these diagnoses. Imaging was semiquantitatively scored and consensus reached using a standard approach by three radiologists blinded to the clinical data. Interestingly, the authors generated several comparator HRCT scores such as “mosaic attenuation or air trapping greater than reticulation” (see later). Multivariable logistic regression identified HRCT features associated with HP and an HRCT-HP diagnosis model was generated using the most highly correlated features. In the analysis, the best performing HRCT variables after controlling for age, sex and smoking status were diffuse axial disease distribution and the novel “mosaic attenuation or air trapping greater than reticulation” score. An HP predictive model using scores based on the binary categorisations of these patterns (*i.e.* present or absent) was created (range 0–3); a score of 3, meaning there is a diffuse axial distribution of disease and mosaic attenuation or air trapping exceeding the extent of reticulation, gave a specificity for HP of >90% in the derivation cohort. The model was then validated in an external cohort generated from the Lung Tissue Research Consortium database (n=438, a score of 3 gave a specificity for HP of 95.8%).

A significant strength of this study is its derivation–validation cohort design, which helps to prevent spurious conclusions being drawn from models that are over-fitted to a single population. However, equally important is clinical applicability. This is especially relevant when diagnostic algorithms are based on HRCT patterns for which interobserver variability is a well-known problem. The authors address this issue head-on by reporting  $\kappa_w$  for each radiologist pair. Interestingly, there was strikingly poor agreement for one of the fundamental patterns in their model, the extent of reticulation ( $\kappa=0.06$  for radiologist 1 *versus* radiologist 3, which is essentially agreement that is no better than by chance). What the authors did not do was to evaluate a rapid scoring of cases based on their final model by a group of radiologists and clinicians of varying levels of experience. Since the model only requires a gestalt impression of the relative extents of mosaic attenuation/air trapping and reticulation, they may have found agreement on the overall model scores a great deal better. This would be a useful test of the algorithms reproducibility.

Clinical applicability aside, there are two difficulties associated with radiological diagnosis that are not addressed. The first is the ubiquitous and often debated issue of diagnostic reference standards. There is no reference standard against which radiological diagnosis can be validated. Histological diagnosis cannot be used because surgical lung biopsy is generally performed in those cases in which HRCT appearances are not definitive, *i.e.* they are often “atypical”; an HRCT study that is based on biopsy diagnosis is a study in which HRCT is intrinsically less helpful and selectively disadvantaged. In addition, histological diagnosis is also subject to high levels of interobserver variability [11], sampling error [12] and may be modified in up to 20% of patients following multidisciplinary discussion [13]. Radiological diagnosis also cannot be standardised against multidisciplinary diagnosis because multidisciplinary diagnosis incorporates and is influenced by radiological diagnosis, and is not, therefore, an independent reference standard [14, 15]. The authors address this issue in the penultimate paragraph of the discussion and emphasise that all attempts to verify clinical diagnosis were blinded to the HRCT findings. For these

reasons, radiological diagnosis should be validated against outcome. This approach has been used in an international study of multidisciplinary IPF diagnosis and to verify the diagnostic accuracy of a novel Deep Learning algorithm for classifying fibrotic lung disease on HRCT [2, 16]. It is also an approach endorsed by the Standards for Reporting of Diagnostic Accuracy initiative [17, 18].

This raises important questions about the utility of our current “histospecific” HRCT diagnostic categories. Although clinicians and patients need to know the “label” of a disease, they also need to know the likely disease behaviour since this will determine clinical management. Aside from patients with usual interstitial pneumonia (UIP) on biopsy, histology does not reliably inform prognosis, and likewise, apart from UIP on HRCT, radiological diagnosis does not reliably predict outcome. Based on the findings of SALISBURY *et al.* [10], a pattern of diffuse axial fibrotic disease with mosaic attenuation and air trapping may allow a confident radiological diagnosis of HP, but it does not tell us about the likely behaviour of the disease. Perhaps, therefore, it is time to rethink radiological diagnosis in ILD; since the primary utility of diagnosis is to inform the clinician of the natural history and treated course of the disease, standardising diagnosis against subsequent disease behaviour may make more sense. Interestingly, these concepts are fuelling a crucial ongoing debate regarding the “progressive fibrotic ILD phenotype”, a label that amalgamates all patients with fibrotic lung disease who exhibit IPF-like disease behaviour regardless of clinical diagnosis. It is an approach formally endorsed in the 2013 ATS/ERS IIP classification statement and argued for in a recently published IPF Working Group statement [14, 19]. For the radiologist, UIP is the classic progressive fibrotic phenotype, but self-sustaining progressive fibrosis is not confined to patients with radiologic UIP. Identifying new radiological phenotypes that reliably inform future disease behaviour (possibly in concert with a blood-based biomarker) may be the future of ILD interpretation on HRCT.

Conflict of interest: None declared.

## References

- 1 Lacasse Y, Girard M, Cormier Y. Recent advances in hypersensitivity pneumonitis. *Chest* 2012; 142: 208–217.
- 2 Walsh SL, Wells AU, Desai SR, *et al.* Multicentre evaluation of multidisciplinary team meeting agreement on diagnosis in diffuse parenchymal lung disease: a case-cohort study. *Lancet Respir Med* 2016; 4: 557–565.
- 3 Johannson KA, Elicker BM, Vittinghoff E, *et al.* A diagnostic model for chronic hypersensitivity pneumonitis. *Thorax* 2016; 71: 951–954.
- 4 Morisset J, Johannson KA, Jones KD, *et al.* Identification of diagnostic criteria for chronic hypersensitivity pneumonitis. An international modified Delphi survey. *Am J Respir Crit Care Med* 2018; 197: 1036–1044.
- 5 Vasakova M, Morell F, Walsh S, *et al.* Hypersensitivity pneumonitis: perspectives in diagnosis and management. *Am J Respir Crit Care Med* 2017; 196: 680–689.
- 6 Lynch DA, Rose CS, Way D, *et al.* Hypersensitivity pneumonitis: sensitivity of high-resolution CT in a population-based study. *AJR Am J Roentgenol* 1992; 159: 469–472.
- 7 Adler BD, Padley SP, Muller NL, *et al.* Chronic hypersensitivity pneumonitis: high-resolution CT and radiographic features in 16 patients. *Radiology* 1992; 185: 91–95.
- 8 Zompatori M, Calabro E, Chetta A, *et al.* [Chronic hypersensitivity pneumonitis or idiopathic pulmonary fibrosis? Diagnostic role of high resolution computed tomography (HRCT)]. *Radiol Med* 2003; 106: 135–146.
- 9 Silva CI, Muller NL, Lynch DA, *et al.* Chronic hypersensitivity pneumonitis: differentiation from idiopathic pulmonary fibrosis and nonspecific interstitial pneumonia by using thin-section CT. *Radiology* 2008; 246: 288–297.
- 10 Salisbury ML, Gross BH, Chughtai A, *et al.* Development and validation of a radiological diagnosis model for hypersensitivity pneumonitis. *Eur Respir J* 2018; 52: 1800443.
- 11 Nicholson AG, Addis BJ, Bharucha H, *et al.* Inter-observer variation between pathologists in diffuse parenchymal lung disease. *Thorax* 2004; 59: 500–505.
- 12 Flaherty KR, Travis WD, Colby TV, *et al.* Histopathologic variability in usual and nonspecific interstitial pneumonias. *Am J Respir Crit Care Med* 2001; 164: 1722–1727.
- 13 Flaherty KR, King TE Jr, Raghu G, *et al.* Idiopathic interstitial pneumonia: what is the effect of a multidisciplinary approach to diagnosis? *Am J Respir Crit Care Med* 2004; 170: 904–910.
- 14 Travis WD, Costabel U, Hansell DM, *et al.* An official American Thoracic Society/European Respiratory Society statement: update of the international multidisciplinary classification of the idiopathic interstitial pneumonias. *Am J Respir Crit Care Med* 2013; 188: 733–748.
- 15 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.
- 16 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; in press.
- 17 Bossuyt PM, Reitsma JB, Bruns DE, *et al.* STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. *Radiology* 2015; 277: 826–832.
- 18 Bruns DE. The STARD initiative and the reporting of studies of diagnostic accuracy. *Clin Chem* 2003; 49: 19–20.
- 19 Wells AU, Brown KK, Flaherty KR, *et al.* What’s in a name? That which we call IPF, by any other name would call the same. *Eur Respir J* 2018; 51: 1800692.