

Title:

Use of Mosaic Attenuation to Distinguish Fibrotic Hypersensitivity Pneumonitis from Idiopathic Pulmonary Fibrosis

Authors:

Barnett J¹, Molyneaux PL², Rawal B¹, Abullah R³, Hare SS³, Vancheesewaran R³, Desai SD¹, Maher TM², Wells AU², Devaraj A¹.

1. Department of Radiology, Royal Brompton Hospital, UK
2. Interstitial Lung Disease Unit, Royal Brompton Hospital, London, UK
3. Royal Free NHS Foundation Trust, London, UK

Abstract:

Background: Mosaic attenuation on CT has been identified in current international guidelines as an important diagnostic feature of fibrotic hypersensitivity pneumonitis (FHP) as opposed to idiopathic pulmonary fibrosis (IPF). However, the term mosaic attenuation (MA) comprises several different radiological signs (low density lobules [LDL], preserved lobules [PL], air trapping [AT] as well as the so-called 'headcheese sign') which may each have different diagnostic utility. Furthermore, the extent of MA required on CT to distinguish these two diagnoses is uncertain, and thresholds of MA from international guidelines have not been validated.

Methods: Inspiratory and expiratory CTs were evaluated by two readers in 102 patients (IPF n=57; FHP n=45) using a semiquantitative scoring system for MA. Findings were validated in an external cohort of patients from a secondary referral institution.

Results: A requirement for increasing extent of LDL and AT based on guidelines (ATS and Fleischner Society) was associated with increased specificity for the diagnosis of FHP (0.96 and 0.98, respectively) at the cost of reduced sensitivity (0.16 and 0.20 respectively). The 'headcheese' sign was found to segregate with high confidence cases of fibrotic HP (positive predictive value 0.85), high specificity of the headcheese sign was maintained in the validation cohort.

Conclusion: The headcheese sign can be confidently considered as being inconsistent with a diagnosis of IPF.

Introduction:

A common dilemma in the diagnosis of fibrosing lung disease is distinguishing idiopathic pulmonary fibrosis (IPF) from fibrotic hypersensitivity pneumonitis (FHP). This distinction is important because these two conditions have different treatments and outcomes.

Several studies have identified the presence of mosaic attenuation (MA) and air trapping (AT) on computed tomography (CT) as being important in the radiological diagnosis of HP^{1,2} and FHP³. This is reflected in the 2011 ATS/ERS/JRS/ALAT guidelines for diagnosis of IPF, which describe diffuse mosaic attenuation/air-trapping (bilateral, in three or more lobes) as being 'Inconsistent with UIP'⁴. Furthermore, a recent update to the ATS/ERS/JRS/ALAT guidelines⁵ and the recent Fleischner Society White Paper⁶ describes '*marked*' or '*extensive*' mosaic attenuation' as a finding 'suggestive of an alternative diagnosis to IPF' or 'consistent with a non IPF diagnosis' respectively, specifically hypersensitivity pneumonitis.

Despite the prominence in diagnostic guidelines of MA and AT as a CT feature suggestive of FHP, and inconsistent or atypical for IPF, these recommendations have not previously been validated, in particular with regard to their ability to distinguish IPF and FHP.

Furthermore, a number of questions arise as to how these CT signs should be used in clinical practice. Firstly, several of the studies which have included MA in

diagnostic models of HP include non-fibrotic HP patients in these cohorts^{1,2}, which may limit the applicability to patients with fibrosis.

Secondly, there is considerable variation in the use of terminology. Although MA is used as a single term in current guidelines^{5,6}, mosaic attenuation consists of a group of different radiological entities, which may have variable diagnostic performance. MA may be characterised by either lobules of low attenuation lung, or alternatively by patchy ground glass opacity adjacent to preserved lobules of normal lung^{7,8}. An additional term which has been referred to as the 'headcheese' sign⁹ represents the combination of preserved lobules and low density lobules within the same lobe, and has recently been identified by an international survey of experts as being an important CT feature of FHP¹⁰. Air trapping (AT), meanwhile is seen exclusively on expiratory CT, and reflects low density lobules seen at end expiration⁷ (figure 1). The diagnostic utility of these various definitions of MA when applying diagnostic guidelines has also, to the best of our knowledge, not been previously assessed.

Therefore the aims of our study were to; 1) investigate the comparative frequency of MA in patients with FHP and IPF based on thresholds from published guidelines, and 2) interpret and validate the various thresholds of MA extent used in guidelines in distinguishing IPF and FHP. For both aims, we examined the impact of using differing definitions of MA, including AT identified on expiratory CT.

Methods

This retrospective study was approved by our institutional review board and patient consent was not required.

Patient selection

143 consecutive patients with an MDT diagnosis of FHP (n=64) or IPF (n=79) undergoing paired inspiratory and expiratory CT at our institution between September 2017 and September 2018 were retrospectively identified from a search of our institutions electronic patient record. Given previous studies questioning the utility of MA in distinguishing FHP and IPF¹¹, little weight is assigned to MA in our institution, when present as an isolated feature, to categorize CTs as inconsistent with UIP. Nevertheless, to reduce the implicit radiological bias of using MDT diagnosis as the gold standard, case review was employed to identify 'high confidence' diagnoses of FHP and IPF (supplementary appendix figure 1) requiring independent evidence of diagnosis. For patients with FHP this included evidence of FHP on lung biopsy (n=14); or a significant exposure history with exposure temporally related to the onset of symptoms (n = 20); or BAL lymphocytosis (>25%) and compatible disease behaviour (steroid responsiveness) (n = 4). In cases where disease behaviour (n=4) or BAL lymphocytosis (n = 3) only were present, CTs were reviewed by an independent radiologist to confirm non-MA signs of FHP as supporting evidence – namely upper or mid zone predominance of fibrosis (n=2), peribronchovascular distribution of disease (n=4) or diffuse ground glass opacity (n=1)⁶. Following

the process of case review, 45 patients with FHP were included in the final derivation cohort.

For patients with IPF high confidence diagnosis were defined as those with an MDT diagnosis of IPF, with a UIP pattern on radiology, age >60 and evidence of progressive disease, in the absence of evidence of connective tissue disease and in the absence of any of the criteria for FHP. Excluded patients were characterised by lack of UIP pattern on CT (n=2), autoimmune features (n=3), exposure history with temporal relation to the onset of symptoms (n=8), atypical histology (n=1), BAL lymphocytosis (n=1) and absence of progressive disease course (n=7). 57 patients with IPF were included in the final derivation cohort.

For the validation cohort, an interstitial lung disease database at a separate, secondary referral institution was used to validate the CT findings, using identical definitions of high confidence disease as in the derivation cohort (supplementary appendix 1, figure 2). The final validation cohort consisted of 65 patients (37 patients with IPF and 28 patients with FHP).

CT Protocol

CTs were acquired with a 128 slice multiple detector CT scanner (Somatom Sensation, Siemens, Erlangen, Germany). Inspiratory CTs were performed at end inspiration at 120Kv with modulation of current, reference 100 mAs. Expiratory CTs were performed at end expiration, following patient training, at 120KV and 30mAs.

In the validation cohort, CTs were acquired with a 128 slice multiple detector CT scanner (Somatom Sensation, Siemens, Erlangen, Germany). Inspiratory CTs

were performed at end inspiration at 120Kv with modulation of current, reference 100 mAs. Expiratory CTs were performed as interspaced studies at end expiration, following patient training, at 120KV with modulation of current, reference 100 mAs.

CT analysis

Mosaic attenuation

Mosaic attenuation was categorised and scored using each of the following three definitions: i) 'Low density lobules', defined as well demarcated regions of low density containing constricted vessels (figure 1). ii) 'Preserved lobules', defined as well demarcated lobules of normal lung surrounded by patchy ground glass opacity (figure 1). iii) 'Air trapping' was assessed on expiratory CT and was defined as well demarcated lobules of low density lung demonstrating a less than normal increase in attenuation, and lack of volume reduction, at expiration⁷ (figure 1)

Each MA variable was scored on a 7-point categorical scale (supplementary appendix figure 3) on a lobar basis by two radiologists (AD, JB) independently, with the lingula defined as a sixth lobe. Post-hoc combination of low density lobules and preserved lobules present in the same lobe was considered representative of the headcheese sign⁹ (figure 2).

Consensus by agreement was performed in cases of disagreement between readers regarding the presence or absence of each variable on a lobar basis. For post-hoc validation of MA scores against RV/TLC and MEF 25%-75%, (see

below), whole lung scores were calculated by summing the lobar low density lobule scores for each patient.

ATS/ERS/JRS/ALAT and Fleischner Society Criteria for CT signs of MA and AT 'inconsistent with UIP' or "suggestive of FHP as an alternative diagnosis to IPF"

The 2011 ATS/ERS/JRS/ALAT guidelines state that the presence of MA/AT in 3 or more lobes bilaterally is 'inconsistent with UIP'⁴. The 2017 Fleischner Society White Paper and the 2018 ATS/ERS/JRS/ALAT guidelines (subsequently referred to as ATS/FS) do not state specific numerical values for MA that indicate FHP as an alternative diagnosis to IPF, but instead use the descriptors 'extensive'⁶ and 'marked'⁵. Hence, we interpreted these terms as representing at least the same extent provided in the 2011 ATS guidance.

We therefore tested the following thresholds: 1) At least 1 lobule of MA in each of three lobes bilaterally (referred to as 'ATS 2011 criteria'). 2) ≥ 3 lobules of MA in each of three or more lobes bilaterally (referred to as 'current ATS/FS criteria, ≥ 3 lobules'), and 3) ≥ 5 lobules of MA in each of three or more lobes bilaterally (referred to as 'current ATS/FS criteria, ≥ 5 lobules').

Fibrosis and Emphysema Extent

Fibrotic lung was defined by honeycomb cysts, reticulation or ground glass opacity causing tractional airways dilatation. Both the fibrosis and emphysema extent were scored to the nearest 5% on a lobar basis by two readers. The mean reader score was calculated, and then the mean of this score across all 6 lobes was taken as the whole lung score.

In the validation cohort, CT data were analysed by two thoracic radiologists (AD and SH) independently for signs of MA, at a threshold extrapolated from the derivation cohort in a binary manner. Fibrosis and emphysema extent were also scored using the derivation cohort methodology.

Pulmonary function tests

96/102 patients (94.1%) in the derivation cohort had pulmonary function test data within 3 months of CT imaging available from institutional records.

Spirometry, diffusion capacity for carbon monoxide (DLco) and plethysmographic lung were performed using established protocols and expressed as a percentage of predicted values¹³. 88/96 (91.7%) patients had maximal expiratory flow measurements available (MEF 25-75%), 73/96 (76.0%) patients were able to complete the flow volume loop for estimation of RV/TLC.

Statistical analysis.

Data and statistical analysis was performed in R (<https://www.R-project.org/>). Interobserver agreement was measured using the interclass correlation coefficient for continuous variables and Cohens kappa with quadratic weighting for categorical data. Comparison of proportions of patients demonstrating MA at various guideline thresholds were made using chi-squared test for proportion or Fisher exact test. The calculation of test statistics was made using 'epiR', confidence intervals for likelihood ratios were performed based on the methods of Simel¹⁴. Linear models were constructed post-hoc to assess the validity of visual scoring metrics against lung function parameters of air trapping and

disease extent, independence of observations and homoscedasticity were confirmed using the Durban Watson test statistic, qq plots and Breush-Pagan test.

Results.

Patient Demographics and Observer agreement

The interclass correlation coefficient for fibrosis and emphysema extent were 0.84 and 0.92. Cohen's kappa with quadratic ratings were good ($k=0.65$ and $k=0.65$) for low density lobules and air trapping and moderate ($k=0.48$) for preserved lobules.

Patient demographics are outlined in Table 1. Patients with IPF were older, more likely to be male, and with a greater pack-year history. Patients with FHP had significantly more extensive fibrosis scored visually, with lower DLco.

Table 1: Patient Characteristics

Characteristic	IPF n=57	FHP n=45	P-value
Demographics			
Men (%)	48 (84.2%)	18 (40%)	<0.001
Age, y ($\pm$ SD)	74.2 ($\pm$ 6.6)	63.2 ($\pm$ 9.3)	<0.001
Ever smoker	39 (69.6%)	18 (40.1%)	0.004
Pack year history (IQR)	17 (31.0)	0 (12.5)	0.005
PFTs*			
DLco ($\pm$ SD)	44.6 ($\pm$ 12.7)	38.0 ($\pm$ 14.7)	0.02
FVC ($\pm$ SD)	74.5 ($\pm$ 19.4)	67.6 ($\pm$ 23.4)	0.1
MEF 25-75 ($\pm$ SD) [§]	82.0 ($\pm$ 38.6)	55.2 ($\pm$ 27.5)	<0.001
RV/TLC ($\pm$ SD) [†]	92.7 ($\pm$ 15.4)	115.0 ($\pm$ 27.9)	<0.001
HRCT			
Fibrosis extent ($\pm$ SD)	19.0 (9.2)	26.3 (16.2)	0.009
Emphysema presence (%)	16 (28.1%)	9 (20%)	0.37

Data are presented as absolute numbers and percentage (sex, smoking status, emphysema presence) and compared with Fischer exact test; mean and standard deviation (age, PFTs, fibrosis extent) and compared with t-test; or as median and interquartile range (pack year history) and compared with the Mann-Whitney test. *PFTs available in 96 patients within 3 months of scan (IPF = 53, FHP = 43) [§] 8 patients with IPF were unable to perform MEF measurements. [†] 11 patients with FHP and 12 patients with IPF were unable to perform RV/TLC measurements. PFT, pulmonary function test; SD, standard deviation; IQR, interquartile range; DLco, diffusing capacity of the lung for carbon monoxide; MEF, maximal expiratory flow; RV/TLC, residual volume / total lung capacity.

Frequency of MA in IPF and FHP according to 2011 ATS and current ATS/FS criteria.

The frequency of patients meeting ATS 2011 and current ATS/FS criteria for FHP as an alternative diagnosis to IPF using different definitions of MA is demonstrated in Table 2.

Table 2. Frequency of CT signs of MA present according to 2011 ATS criteria or current ATS/FS thresholds for FHP as an alternative diagnosis to IPF

MA Variable	IPF n=57	FHP n=45	P-value
Preserved lobules			
ATS 2011 criteria	11 (19.3%)	27 (60%)	<0.001
Current ATS/FS, ≥ 3 lobules	4 (7.0%)	15 (33.3%)	<0.001
Current ATS/FS, ≥ 5 lobules	1 (1.8%)	8 (17.8%)	0.009
Low density lobules			
ATS 2011 criteria	29 (50.9%)	35 (77.8%)	0.005
Current ATS/FS, ≥ 3 lobules	8 (14.0%)	23 (51.1%)	<0.001
Current ATS/FS ≥ 5 lobules	2 (3.5%)	7 (15.6%)	0.04
Headcheese Sign			
ATS 2011 criteria	4 (7.0%)	22 (48.9%)	<0.001
Current ATS/FS, ≥ 3 lobules	4 (7.0%)	22 (48.9%)	<0.001
Current ATS/FS, ≥ 5 lobules	2 (3.5%)	19 (38.8%)	<0.001
Expiratory Air Trapping			
ATS 2011 criteria	26 (45.6%)	37 (82.2%)	<0.001
Current ATS/FS, ≥ 3 lobules	5 (8.8%)	22 (48.9%)	<0.001
Current ATS/FS, ≥ 5 lobules	1 (1.8%)	9 (20%)	0.004

Frequency of various CT signs of MA by diagnosis, using different definitions of MA and different thresholds of MA for FHP based on guidelines. 'ATS 2011 criteria' require at least one lobule in 3 or more lobes, bilaterally. 'Current ATS/FS ≥ 3 lobules' requires 3 or more lobules in 3 or more lobes bilaterally, 'Current ATS/FS, ≥ 5 lobules' requires 5 or more lobules in 3 or more lobes bilaterally. Headcheese represents the sum of both low density lobules and preserved lobules only when both are present in the same lobe. MA; mosaic attenuation, ATS; American Thoracic Society, FS; Fleischner Society.

Diagnostic utility of MA for diagnosis of FHP

Table 3 demonstrates the test statistics for the various definitions and thresholds of MA; a positive test result is a patient meeting the threshold of FHP as an alternate diagnosis to IPF using varying thresholds and definitions for MA; a true positive test is a patient with a high confidence (incorporating MDT diagnosis and case review) diagnosis of fibrotic hypersensitivity pneumonitis.

The presence of the headcheese sign in three or more lobes, bilaterally, was highly specific (0.93) for a diagnosis of FHP, and moderately sensitive (0.49). On the contrary, expiratory air trapping, and its counterpart on inspiratory images, low density lobules, were sensitive (0.82 and 0.78 respectively) but only moderately specific for FHP, (0.54 and 0.49), when interpreted at the 2011 thresholds.

Applying more stringent thresholds of MA to suggest FHP as an alternate diagnosis to IPF, ('Current ATS/FS $\geq$ 3 lobules' and 'Current ATS/FS $\geq$ 5 lobules'), showed increased specificity of low density lobules (0.86 and 0.96 respectively), and expiratory air trapping (0.91 and 0.98 respectively). These more stringent thresholds ('Current ATS/FS $\geq$ 3 lobules' and 'Current ATS/FS $\geq$ 5 lobules'), were associated with reduced sensitivity, relative to 2011 ATS criteria, for low density lobules (0.51, 0.16 respectively) and expiratory air trapping (0.49, 0.2 respectively).

Table 3. Test statistics for 2011 ATS and extrapolated current ATS/FS thresholds for FHP as an alternative diagnosis to IPF.

MA Variable	Sensitivity	Specificity	PPV	NPV	+LR	-LR
Preserved lobules						
ATS 2011 criteria	0.60	0.81	0.71	0.72	3.11 (1.74, 5.56)	0.50 (0.34, 0.72)
Current ATS/FS, ≥ 3 lobules	0.33	0.93	0.79	0.64	4.75 (1.69, 13.32)	0.72 (0.58, 0.89)
Current ATS/FS, ≥ 5 lobules	0.18	0.98	0.89	0.60	10.13 (1.32, 78.07)	0.84 (0.73, 0.96)
Low density lobules						
ATS 2011 criteria	0.78	0.49	0.55	0.74	1.53 (1.13, 2.06)	0.45 (0.25, 0.83)
Current ATS/FS, ≥ 3 lobules	0.51	0.86	0.74	0.69	3.64 (1.80, 7.36)	0.57 (0.41, 0.78)
Current ATS/FS, ≥ 5 lobules	0.16	0.96	0.78	0.59	4.43 (0.97, 20.31)	0.88 (0.76, 1.00)
Headcheese						
ATS 2011 criteria	0.49	0.93	0.85	0.7	6.97 (2.59, 18.77)	0.55 (0.41, 0.74)
Current ATS/FS, ≥ 3 lobules	0.49	0.93	0.85	0.7	6.97 (2.59, 18.77)	0.55 (0.41, 0.74)
Current ATS/FS, ≥ 5 lobules	0.42	0.96	0.9	0.68	12.03 (2.96, 48.97)	0.60 (0.46, 0.77)
Expiratory Air Trapping						
ATS 2011 criteria	0.82	0.54	0.59	0.79	1.8 (1.32, 2.47)	0.33 (0.17, 0.64)
Current ATS/FS, ≥ 3 lobules	0.49	0.91	0.81	0.69	5.57 (2.29, 13.56)	0.56 (0.42, 0.75)
Current ATS/FS, ≥ 5 lobules	0.2	0.98	0.9	0.61	11.4 (1.50, 86.69)	0.81 (0.70, 0.95)

Test statistics for the MA variables using various thresholds; 2011 ATS criteria and extrapolated 2018 ATS and Fleischner Society White Paper ('Current ATS/FS') criteria. Meeting the threshold for FHP as an alternate diagnosis to IPF is denoted as a positive test result. A true positive is a patient with FHP. Confidence intervals

expressed in brackets for the positive (+LR) and negative (-LR) likelihood ratios. PPV; positive predictive value, NPV; negative predictive value.

Validation of visual scoring metrics against lung function

To confirm that MA identified on inspiratory CTs in this study represented physiological air trapping, post hoc linear models were created to examine the relationship of low density lobule scores with pulmonary function indices of air trapping; RV/TLC (n = 73) and MEF 25-75% (n = 88). Low density lobule scores were found to be significantly correlated with both air trapping indices, after adjustment for visual fibrosis and emphysema (table 4, 5). In order to assess for disease extent confounding our visual scores of air trapping, a separate linear model found that both visual fibrosis and emphysema scores were correlated with DLco, however visual low density lobule score was not (table 6).

Table 4: Linear model of RV/TLC against visual scores of whole lung low density lobule, fibrosis and emphysema extent.

	β Estimate	Std Error	P value
Low density lobule	2.00	0.42	<0.001
Fibrosis	0.28	0.20	0.16
Emphysema	0.08	0.25	0.73

Adjusted R-squared: 0.29, p-value: <0.001

Table 5: Linear model of MEF 25-75% against visual low density lobule score, visual fibrosis extent and visual emphysema extent.

	β Estimate	Std Error	P value
Low density lobule	-2.54	0.64	<0.001
Fibrosis	-0.13	0.28	0.66
Emphysema	-0.77	0.41	0.06

Adjusted R-squared: 0.15, p-value: 0.001

Table 6: Linear model of DLco against visual low density lobule score, visual fibrosis extent and visual emphysema extent.

	β Estimate	Std Error	P value
Low density lobule	-0.36	0.22	0.1
Fibrosis	-0.56	0.09	<0.001
Emphysema	-0.46	0.13	<0.001

Adjusted R-squared: 0.33, p-value: <0.001

Validation of findings against an external cohort

In order to validate our findings, the threshold of ‘three of more lobules, in three or more lobes, bilaterally’ was selected as the optimal balance between high specificity and acceptable sensitivity for a diagnosis of FHP, as an alternate diagnosis to IPF. This threshold was applied to CT data from an external cohort, derived from a database of interstitial lung disease at a secondary referral institution for interstitial lung disease (supplementary figure 2). Agreement between the two scorers in the validation cohort was substantial for low density lobules (k=0.61) and headcheese (k=0.71) and moderate for expiratory air trapping (k=0.56). As in the derivation cohort, patients with IPF were older, more frequently male, more frequently ever smokers and had a greater smoking history (table 7).

Table 7, Patient characteristics, validation cohort

Characteristic	IPF n=37	FHP n=25	P-value
Demographics			
Men* (%)	34 (91.9%)	12 (48%)	0.001
Age, y ($\pm$ SD)	76.3 (9.3)	71.6 (8.9)	0.05
Ever smoker	24 (64.9%)	8 (32%)	0.02
Pack year history (IQR)	20.0 (39)	0.0 (3)	0.004
HRCT			
Fibrosis extent ($\pm$ SD)	33.1 (12.7)	23.7 (14.6)	0.02
Emphysema presence (%)	10 (27.0%)	4 (16.0%)	0.4

*Data are presented as absolute numbers and percentage (sex, smoking status, emphysema presence) and compared with Fischer exact test; mean and standard deviation (age, fibrosis extent) and compared with t-test; or as median and interquartile range (pack year history) and compared with the Mann-Whitney test. SD, standard deviation; IQR, interquartile range; ***Sex defined by gender at birth.***

Test statistics for the validation cohort are displayed in table 8. The headcheese sign retained its high specificity (0.92,0.95) for both readers, whereas the sensitivity was greater than that seen in the derivation cohort (0.76, 0.64). The specificity of low density lobules (0.65, 0.65) and expiratory air trapping (0.76, 0.65) were however decreased relative to the derivation cohort.

Table 8, Test statistics for visual scores of different MA variables in the validation cohort.

	Reader 1		Reader 2	
	Sensitivity	Specificity	Sensitivity	Specificity
Low density lobules				
Current ATS/FS, ≥ 3 lobules	0.72	0.65	0.8	0.65
Headcheese Sign				
Current ATS/FS, ≥ 3 lobules	0.76	0.92	0.64	0.95
Expiratory Air Trapping				
Current ATS/FS, ≥ 3 lobules	0.64	0.76	0.80	0.65

Test statistics for the MA variables using the ‘three or more lobules, in three or more lobes, bilaterally’ threshold extrapolated from 2018 ATS and Fleischner Society White Paper criteria. Meeting the threshold for FHP as an alternate diagnosis to IPF is denoted as a positive test result. A true positive is a patient with FHP.

Discussion

In our study, we have attempted to validate current diagnostic guidelines which are in widespread use by radiologists and pulmonologists when attempting to differentiate FHP from IPF on CT. We have shown that MA is a common finding in IPF, being present in up to 45% of patients at the preceding ATS guidelines, in line with two previous studies^{3,11} and indicate that MA should not be considered as a feature that is necessarily inconsistent with, or atypical for IPF.

This point is reinforced when evaluating the guidance described in current diagnostic guidelines. In the absence of a specific definition of the terms ‘marked’⁵ and ‘extensive’⁶ used in current guidelines, we chose to extrapolate the

2011 ATS/ERS/JRS/ALAT criteria as this had a specific, reproducible threshold. We chose two interpretations of varying extent; requiring ≥ 3 lobules, and ≥ 5 lobules of each MA variable, in three or more lobes, bilaterally. We found that although increasing the required extent of MA on CT to suggest FHP as an alternative diagnosis to IPF improved diagnostic specificity, even at the most extensive extrapolation of the Fleischner Society White Paper⁶ and 2018 ATS/ERS/JRS/ALAT clinical practice guidelines⁵ (requiring 5 lobules of a given MA variable within 3 or more lobes, bilaterally), no threshold of MA could completely exclude a diagnosis of IPF. Furthermore, such a profuse extent of MA was associated with poor sensitivity, whether defined by preserved lobules, low density lobules or expiratory air trapping, thus potentially limiting the clinical utility of these CT features.

The 'headcheese' sign, first described by Webb⁹, is common parlance amongst radiologists and refers to a specific form of mosaic attenuation, describing the combination of ground glass opacity, preserved lobules and low attenuation lobules. Indeed, in a recent Dephi survey¹⁰, this sign reached consensus as important in the diagnosis of fibrotic hypersensitivity pneumonitis. The headcheese sign has, however, not been previously examined with regard to its diagnostic utility. We found the presence of the 'headcheese sign' to be strongly specific for a diagnosis of FHP, outperforming all other MA metrics, and thus out of all the MA variables can be considered atypical for IPF and more suggestive of FHP, at any threshold. Indeed, the headcheese sign retained high specificity when examined by two independent radiologists in an external cohort of patients with IPF and FHP.

Our findings show that the headcheese sign co-segregates with high confidence clinical features of fibrotic HP, and is rarely associated with high confidence clinical variables of IPF. Whilst it would not be robust to extrapolate the utility of the headcheese sign to low confidence or unclassifiable diagnoses (which were not examined in this cohort), the headcheese sign can be regarded as significant when the MDT is faced with incomplete clinical data. Furthermore, the headcheese sign is widely known amongst radiologists; and has the potential to be a distinctive, widely useable CT feature suggestive of FHP as an alternative diagnosis to IPF.

When applied to the validation cohort, the high specificity of the Current ATS/FS, ≥ 3 lobules threshold was not maintained for low density lobules or air trapping. This combination of decreased specificity and increased sensitivity may reflect increased incidence of air trapping in this population. The variability of these findings limits the confidence of low density lobules and expiratory air trapping at the thresholds examined, when seen in the absence of the headcheese sign.

MA has previously been described as an important CT parameter in diagnostic models of hypersensitivity pneumonitis^{1,2}. However, one important distinction which should be made between this study and previous diagnostic models of HP^{1,2}, is that our cohort contains only patients with fibrotic (as opposed to non-fibrotic) HP. Patients with non-fibrotic HP, by definition, fall within a different diagnostic conundrum where IPF is not a valid differential diagnosis. The

applicability of diagnostic models derived from partly non-fibrotic cohorts to distinguish fibrotic HP from IPF is uncertain.

The rationale for employing CT features of MA as a sign of FHP is based, in part, on the airways centric fibrosis demonstrated on histopathology. Whilst a UIP histology lacks implicit airways disease, a major confounder is smoking history, which is often encountered in patients with IPF¹⁵. A previous radiological study in 250 healthy subjects found 27% of ex-smokers to demonstrate expiratory air trapping of segmental or lobar extent, a significantly more frequent finding than in never smokers¹⁶. This may explain the high frequency of MA and AT in patients with IPF in our and previous studies¹¹.

The examination of visual scores against RV/TLC, MEF_{25-75%} and DLco was undertaken to examine the reliability of our visual scores of air trapping, and the possible confounding of emphysema on both visual scores and pulmonary function tests of small airways disease. Emphysema effects both measures of airflow and DLco, whereas intrinsic airways disease should not influence DLco¹⁷, unless severe. The association of our visual scores of air trapping with both airflow metrics, and their independence from DLco; suggest our visual scores are robust.

One limitation to our study is the reliance of MDT diagnosis to establish gold-standard diagnosis. Radiologist interpretation of CT images forms a key part of the MDT reaching a diagnosis, potentially confounding any study on CT findings. We believe our methodology is robust for two reasons. Firstly, since the findings

of the IPFnet study¹¹, we have given little weight to low density lobules suggestive of physiological air trapping in our MDT, especially when present in a patient who would be otherwise of 'definite' or 'probable' UIP morphology. Secondly, our case review ensured 'high-confidence' diagnoses only were included, excluding cases based solely on radiological findings. Whilst the criteria used for this case review have been shown to be widely accepted in the diagnosis of fibrotic hypersensitivity pneumonitis¹⁰, the lack of standardised international criteria for diagnosis of FHP adds heterogeneity to this cohort. Our institution acts as a tertiary referral centre for ILD, this may introduce patient selection bias towards more severe disease. However, our findings with respect to the headcheese sign were also applicable in a secondary referral centre population.

In conclusion, we have shown that the individual components of the heterogeneous term of 'mosaic attenuation' have variable diagnostic performance, and therefore caution should be applied when referring to the term 'mosaic attenuation' without clarification. Thresholds for 'extensive' or 'marked' mosaic attenuation based on current guidelines are associated with improved specificity but reduced sensitivity for FHP. The headcheese sign, however, should be confidently considered a CT sign of FHP as an alternate diagnosis to IPF.

References

1. Johansson KA, Elicker BM, Vittinghoff E, et al. A diagnostic model for chronic hypersensitivity pneumonitis. *Thorax*. 2016;71(10):951-954. doi:10.1136/thoraxjnl-2016-208286
2. Salisbury ML, Gross BH, Chughtai A, et al. Development and validation of a radiological diagnosis model for hypersensitivity pneumonitis. *Eur Respir J*. 2018;52(2):1800443. doi:10.1183/13993003.00443-2018
3. Silva CIS, 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(1):288-297. doi:10.1148/radiol.2453061881
4. 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(6):788-824. doi:10.1164/rccm.2009-040GL
5. 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(5):e44-e68. doi:10.1164/rccm.201807-1255ST
6. Lynch DA, Sverzellati N, Travis WD, et al. Diagnostic criteria for idiopathic pulmonary fibrosis: A Fleischner Society White Paper. *Lancet Respir Med*. 2017;2600(17):1-16. doi:10.1016/S2213-2600(17)30433-2
7. Hansell DM, Bankier AA, MacMahon H, McLoud TC, Müller NL, Remy J. Fleischner Society: glossary of terms for thoracic imaging. *Radiology*. 2008;246(3):697-722. doi:10.1148/radiol.2462070712
8. Stern EJ, Swensen SJ, Hartman TE, Frank MS. CT mosaic pattern of lung attenuation: distinguishing different causes. *Am J Roentgenol*. 1995;165(4):813-816. doi:10.2214/ajr.165.4.7676972
9. Webb WR (Wayne R, Müller NL, Naidich DP. *High-Resolution CT of the Lung*. Wolters Kluwer/Lippincott Williams & Wilkins; 2009.
10. Morisset J, Johansson 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(8):1036-1044. doi:10.1164/rccm.201710-1986OC
11. Yagihashi K, Huckleberry J, Colby T V., et al. Radiologic-pathologic discordance in biopsy-proven usual interstitial pneumonia. *Eur Respir J*. 2016;47(4):1189-1197. doi:10.1183/13993003.01680-2015
12. Ryerson CJ, Corte TJ, Lee JS, et al. A standardized diagnostic ontology for fibrotic interstitial lung disease an international working group perspective. *Am J Respir Crit Care Med*. 2017;196(10):1249-1254. doi:10.1164/rccm.201702-0400PP
13. Quarijer PH, Tammeling GJ, Cdl JE, Pedersen OF, Peslin R, Yernault J-C. Lung Volumes and Forced Ventilatory Flows. *Eur Respir J*. 1993;6:5-40.
14. Simel DL, Samsa GP, Matchar DB. Likelihood ratios with confidence: sample size estimation for diagnostic test studies. *J Clin Epidemiol*. 1991;44(8):763-770.
15. Baumgartner KB, Samet JM, Stidley CA, Colby T V, Waldron JA. Cigarette

- smoking: a risk factor for idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med*. 1997;155(1):242-248. doi:10.1164/ajrccm.155.1.9001319
16. Mastora I, Remy-Jardin M, Sobaszek A, Boulenguez C, Remy J, Edme J-L. Thin-Section CT Finding in 250 Volunteers: Assessment of the Relationship of CT Findings with Smoking History and Pulmonary Function Test Results. *Radiology*. 2001;218(3):695-702. doi:10.1148/radiology.218.3.r01mr08695
 17. Hansell DM, Rubens MB, Padley SP, Wells AU. Obliterative bronchiolitis: individual CT signs of small airways disease and functional correlation. *Radiology*. 1997;203(3):721-726. doi:10.1148/radiology.203.3.9169694