ERJ Express. Published on June 25, 2018 as doi: 10.1183/13993003.00443-2018

EUROPEAN RESPIRATORY *journal*

FLAGSHIP SCIENTIFIC JOURNAL OF ERS

Early View

Original article

Development and Validation of a Radiologic Diagnosis Model for Hypersensitivity Pneumonitis

Margaret L. Salisbury, Barry H. Gross, Aamer Chughtai, Mohamed Sayyouh, Ella A. Kazerooni, Brian B. Bartholmai, Meng Xia, Susan Murray, Jeffrey L Myers, Amir Lagstein, Kristine E Konopka, Elizabeth A. Belloli, Jamie S. Sheth, Eric S. White, Colin Holtze, Fernando J. Martinez, Kevin R. Flaherty

Please cite this article as: Salisbury ML, Gross BH, Chughtai A, *et al*. Development and Validation of a Radiologic Diagnosis Model for Hypersensitivity Pneumonitis. *Eur Respir J* 2018; in press (https://doi.org/10.1183/13993003.00443-2018).

TITLE: Development and Validation of a Radiologic Diagnosis Model for Hypersensitivity Pneumonitis

AUTHORS: Margaret L. Salisbury[1], Barry H. Gross[2], Aamer Chughtai[2], Mohamed Sayyouh[2], Ella A. Kazerooni[2], Brian B. Bartholmai[3], Meng Xia[4], Susan Murray[4], Jeffrey L Myers[5], Amir Lagstein[5], Kristine E Konopka[5], Elizabeth A. Belloli[1], Jamie S. Sheth[1], Eric S. White[1], Colin Holtze[1], Fernando J. Martinez[6], Kevin R. Flaherty[1]

[1] University of Michigan, Division of Pulmonary and Critical Care, Ann Arbor, MI
[2] University of Michigan, Department of Radiology, Ann Arbor, MI
[3] Mayo Clinic, Department of Radiology, Rochester, MN
[4] University of Michigan, Department of Biostatistics, Ann Arbor, MI
[5] University of Michigan, Department of Pathology, Ann Arbor, MI
[6] Cornell Medical College Division of Pulmonary and Critical Medicine, New York, NY, United States of America

Corresponding Author: Margaret L. Salisbury. 1500 East Medical Center Drive, 3916 Taubman Center, Ann Arbor, MI 48109. Phone: (734) 647-6477, Fax: (734) 936-5048, e-mail: msalisbu@med.umich.edu.

Take-Home Summary: When HRCT shows more mosaic attenuation than reticulation and diffuse axial ILD, false HP diagnosis risk is <10%.

Author Contributions: M.L.S., K.R.F., F.J.M. and E.A.K. conceived and designed the study; M.L.S. and M.X. analyzed the data with supervision and assistance from S.K.M. K.R.F., B.J.B., E.A.K., F.J.M., B.H.G., A.C., M.S., J.L.M., A.L., K.E.K., E.A.B., J.S.S., E.S.W., and C.H. contributed data; M.L.S. prepared the manuscript; all authors critically revised the manuscript for intellectual content, approved the final draft, and agree to accountability for all aspects of the work.

Funding Support: T32 HL00749-21 (Multidisciplinary Training Program in Lung Disease), National Institutes of Health K24 HL111316 (Kevin R. Flaherty), and National Institutes of Health/National Heart, Lung, and Blood Institute HHSN26820118C (Lung Tissue Research Consortium).

Subject Code: 9.23 Interstitial Lung Disease

This article has an online data supplement.

ABSTRACT

Rationale: High resolution computed tomography may be useful for diagnosing hypersensitivity pneumonitis (HP). We develop and validate a radiologic diagnosis model and model-based points score.

Methods: Patients with interstitial lung disease seen at University of Michigan (derivation), or enrolling in the Lung Tissue Research Consortium (validation) were included. A thin-section, inspiratory CT was required. Thoracic radiologists documented radiologic features.

Results: The derivation cohort comprised 356 subjects (33.9% HP) and validation cohort 438 (15.2% HP) An age-, gender-, and smoking status-adjusted logistic regression model identified extent of mosaic attenuation or air trapping greater than that of reticulation ("MA-AT>Reticulation"; OR 6.20, CI 95% 3.53-10.90, p<0.0001) and diffuse axial disease distribution (OR 2.33, CI 95% 1.31-4.16, p=0.004) as HP predictors (AUC=0.814). A model-based score >2 (1 point for axial distribution, 2 for MA-AT>Reticulation) has specificity 90% and PPV 74% in the derivation and specificity 96% and PPV 44% in the validation cohort. Similar model performance is seen with population restriction to those reporting no exposure (score>2 specificity 91%).

Conclusions: When radiologic MA or AT are more extensive than reticulation and disease has diffuse axial distribution, HP specificity is high and false diagnosis risk low (<10%), but PPV is diminished in a low-prevalence setting.

Key Words:

Lung Diseases, Interstitial

Multidetector Row Computed Tomography

Pulmonary Fibrosis

INTRODUCTION

Hypersensitivity pneumonitis (HP) is an interstitial lung disease (ILD) caused by exposure to a variety of antigens.(1) High-resolution computed tomography (HRCT) is commonly used in the diagnostic evaluation for ILDs. A specific HRCT pattern in the correct clinical context can be diagnostic of histopathologic usual interstitial pneumonia (UIP), pathognomonic of idiopathic pulmonary fibrosis (IPF) in the correct clinical context.(2) Individuals without a specific ILD diagnosis after clinical and HRCT evaluation are often subjected to additional invasive diagnostic testing (e.g. bronchoscopy or surgical lung biopsy) to obtain a diagnosis.

Several studies have evaluated HRCT patterns associated with HP. Two studies found that a radiologist's confident diagnosis of HP is correct 88-92% of the time.(3, 4) Johannson and colleagues created a clinical prediction model combining patient age, environmental exposure history, and radiologic features of diffusely distributed ground glass or mosaic attenuation, and found that a high model-based score was associated with a high specificity for HP.(5) Unfortunately, these studies included limited alternative ILD diagnoses,(3, 4) or had the potential for confirmation bias by modeling exposure history along with radiologic features.(5) We sought to derive and externally validate a "rule-in" diagnostic model for HP based solely on radiologic findings.

METHODS

Patient Selection and Clinical Diagnosis Assignment

The derivation cohort consisted of a retrospectively-assembled cohort of consecutive adult patients undergoing diagnostic case review at the University of Michigan's (UMHS) multidisciplinary ILD conference between 1-February-2009 and 31-August-2014. Additional HP cases were identified by "hypersensitivity pneumonitis" ICD-9 code (495.7, 495.8, 495.9) electronic medical record search between 1-January-2004 to 31-December-2013. Clinical ILD

diagnoses (the outcome variable) were verified after a detailed chart review. Supporting evidence sufficient for HP diagnosis verification included classic findings on surgical lung biopsy, or at least two of: 1) BAL lymphocytosis >20%;(6, 7) 2) consistent findings on transbronchial or surgical biopsy (any of loose non-necrotizing granulomas, giant cells, mononuclear inflammatory interstitial or peribronchiolar infiltrate); 3) a plausible exposure history.(1) Patients with suspected drug-induced pulmonary hypersensitivity reaction were excluded. Non-HP ILD diagnoses were verified in the setting of consistent histopathologic evaluation and according to current guidelines.(2, 8) Documentation of ILD conference discussion(9) was available in 69 (57%) of 122 identified HP patients and 226 (94%) of 240 identified subjects with non-HP ILDs. Radiology (HRCT) reports were available to the multidisciplinary team and treating clinicians, but were not considered as part of the clinical diagnosis verification process for this study. All subjects had a baseline HRCT, and unclassifiable ILD were excluded. **Figure 1** gives a flow diagram with numbers of screened patients from each dataset, reasons for exclusion, and final included clinical diagnoses for the UMHS derivation cohort. Demographics, clinical characteristics, and pulmonary function measurements were collected from the electronic medical record.

The validation cohort consisted of patients with ILD undergoing clinically-indicated lung tissue sampling procedures as part of the multicenter, NIH-supported Lung Tissue Research Consortium (LTRC) study (www.ltrcpublic.com, last accessed 27 September 2017). Availability of a histopathologic specimen obtained via diagnostic surgical lung biopsy, and complete visual HRCT features scores documentation, were required for inclusion. Baseline demographic and pulmonary function data was available for all patients. Methods of patient collection and assignment of clinical ILD diagnoses for the LTRC study is described in detail elsewhere.(10) **Figure 2** gives a flow diagram with numbers of included and excluded patients, and final included clinical diagnoses, for the LTRC validation cohort.

The Institutional Review Board at the University of Michigan approved this study (HUM 00093978). The LTRC study was approved by participating centers, and approved use de-identified data for this analysis. Some of these results have been presented as an abstract.(11)

Derivation Cohort Radiologic Protocols

Images were obtained on a variety of CT scanners at UMHS and at referring centers, and viewed on a PACS system. All HRCTs had a non-contrasted image series with thin (2mm collimation or less) sections obtained at end-inspiration. Prone and expiratory images and a volumetric or image-sharpening protocol were available in most subjects, but not required for inclusion. HRCT scan quality was documented as excellent, diagnostic, or non-diagnostic. Slice thickness, inter-slice interval (when appropriate), availability of a volumetric protocol, expiratory and prone series were documented. Subjects with a non-diagnostic HRCT or inspiratory-series slice thickness >2mm were excluded (see **Figure 1**).

Three radiologists (B.H.G., A.C., M.S.) blinded to clinical data independently interpreted each HRCT and documented findings on a standardized scoring sheet (see **Figure e1**). Using standard definitions(12), radiologic features of honeycombing, reticular pattern, ground glass, mosaic attenuation, and air trapping (when expiratory images were available) were scored semi-quantitatively in each lobe (right upper lobe, right middle lobe, right lower lobe, left upper lobe/lingula, and left lower lobe), with 0 indicating no involvement, 1 indicating <5% of lobe involved (present but minimal), 2 indicating 5-25%, 3 indicating 25-49%, 4 indicating 50-75%, and 5 indicating >75% involvement. For each feature, the lobe scores were summed and divided by 5 to obtain an average on a scale of 0 to 5, representative of the proportion of total lung having the feature. Traction bronchiectasis and centrilobular nodules were recorded as present (score=1) or absent (score=0) in each of 5 lobes; lobar scores were summed to represent the number of lobes having the feature (scale 0 to 5). A dichotomous variable was created for each feature, considered “present” if the average or sum was greater than 0.5. We selected this cut-point to avoid labeling a radiologic feature as present when its extent was

minimal. For example, a ground glass average of 0.4 means 2 lobes of the lung had ground glass present in 5% or less of the lobe ([1+1]/5=0.4), or 1 lobe had ground glass in 5-25% of the lobe (2/5=0.4). The predominant distribution of interstitial disease was assigned in the axial and craniocaudal dimensions. Axial distribution was noted as central (predominant parenchymal abnormality preferentially involves the central 1/3 of the lung), peripheral (abnormality preferentially involves the peripheral 1/3 of the lung), subpleural (abnormality preferentially involves the immediate subpleural region), peribronchovascular (abnormality preferentially involves the region adjacent to the peribronchovascular bundles), or diffuse (abnormality is widely distributed with no section involved more than another). Multiple selections were allowed for axial distribution but no reader made more than 2 selections. To be consistent with the methodology of the LTRC radiologist (see Validation Cohort Radiologic Protocol Below), a selection of central, peripheral, or diffuse disease was considered the primary distribution when multiple selections were made. When subpleural was the only selection, it was grouped with peripheral. The craniocaudal distribution of disease was noted as upper lung (predominant parenchymal abnormality is most prominent above the main carina), lower lung (abnormality is most prominent below the main carina), or diffuse (upper and lower lungs are relatively equally involved); only one selection was allowed.

Three-reader consensus dichotomous and semi-quantitative HRCT scores were created for each subject. Consensus semi-quantitative scores were a 3-reader average of the average/summed feature score. Dichotomized consensus variables were “present” when 2 of the 3 readers’ scores met “present” criteria. Several additional dichotomized variables were created: Mosaic attenuation or air trapping (“MA-AT”) was “present” when a subject had consensus presence of mosaic attenuation or air trapping. The “GG>Reticulation” variable was present when the semi-quantitative ground glass score was higher than the semi-quantitative reticulation score. The “MA-AT>Reticulation” variable was present when the semi-quantitative score for mosaic attenuation or air trapping was higher than the semi-quantitative reticulation

score. Consensus for the craniocaudal and axial disease distribution was the probability of a radiologist assigning the distribution category for each patient. A consensus dichotomous variable for a given category (i.e. upper, lower, or diffuse) of the disease distribution was assigned as “present” when its probability was 2/3 or greater.

Validation Cohort Radiologic Protocols

The LTRC enrollment HRCT was assessed by one expert thoracic radiologist in the LTRC Radiology Core Laboratory. The semi-quantitative scoring procedure and a list of radiologic features analyzed is described in detail elsewhere.(10) A dichotomous variable and a semi-quantitative average (representative of the proportion of total lung having the feature) were created for the HRCT features of honeycombing, reticulation, ground glass, mosaic attenuation, air trapping, traction bronchiectasis, and centrilobular nodules. The scale for semi-quantitative average scores is 0-4 and represents quartiles of lung involvement with the feature (e.g. a score of 1 is 1-25%, and 4 is 76-100% of lung). A feature was dichotomously “present” when the semi-quantitative average was greater than 0.5. Axial (none, diffuse/even, central, peripheral/subpleural) and craniocaudal (upper, lower, even/diffuse) disease distribution were also available.

Statistical Methods

Analyses were performed using SAS version 9.4 (SAS Institute, 2013) and R version 3.2.1 (The R Foundation, 2015). Baseline patient and HRCT quality characteristics are shown as mean (SD), or number (percentage), as appropriate. Highly skewed variables are shown as median (range). Pairwise and 3-way agreement of 3 radiologists on the dichotomized HRCT features are via the kappa and Light’s kappa, respectively.(13) The bootstrap method was used to calculate 95% confidence interval for Light’s kappa. Pairwise agreement on semi-quantitative features is via weighted kappa, with the averaged or summed feature score (scale 0-5) rounded to the nearest integer. Kappa (κ) measures indicate that agreement is poor when $\kappa<0.40$, intermediate when $0.40<\kappa<0.60$, good when $0.60<\kappa<0.75$, and excellent when $0.75<\kappa$.(14) To

identify the nature of systematic differences in interpretation of continuous radiologic features across the 3 readers, we also evaluated the mean (SD) of each feature for each pair of radiologists using paired t-tests. For dichotomous features, the number (%) of patients having the feature “present” by radiologists is shown, with significance of pairwise between-reader differences via McNemar’s test.(15)

Logistic regression models identified clinical and radiologic variables associated with HP. Area under the receiver-operating curve (AUC) gives model discrimination.(16) An HP-HRCT Diagnosis Score was created from the final multivariable model. To validate the model, we obtained the predicted probability of HP from both the final regression formula developed in the UMHS derivation cohort, and from the HP-HRCT Diagnosis Score, for each subject in the LTRC dataset. HP-HRCT Diagnosis Score test characteristics are given as sensitivity and specificity at various points thresholds. To measure model calibration, LTRC subjects were categorized into discretized groups of predicted model-based probabilities of HP diagnosis (using the UMHS final full model and adjusted HP-HRCT diagnosis score models) and then compared to observed frequencies of HP diagnosis within each probability group. The resulting observed versus expected probabilities are displayed graphically with a superimposed line of perfect calibration shown for reference.(17) We also performed several sub-population sensitivity analyses within the UMHS cohort: (1) with HP subjects restricted to include only those with “classic” lung biopsy findings; (2) with HP subjects restricted to “classic” lung biopsy findings and non-HP subjects restricted to idiopathic interstitial pneumonias only (acute interstitial pneumonitis, cryptogenic organizing pneumonia, desquamative interstitial pneumonia, idiopathic pulmonary fibrosis, nonspecific interstitial pneumonitis, and respiratory bronchiolitis-ILD); and (3) with all subjects restricted to those with no documented exposure history.

RESULTS

Patient Characteristics

In the UMHS cohort, the HP group was slightly younger, had a higher proportion of female patients, a higher proportion of never-smokers, and a higher proportion were treated with corticosteroids prior to the consultation at UMHS (**Table 1)**. In the LTRC cohort, the HP group was slightly younger, had a greater proportion of never-smokers, and had a higher DLCO% than the not HP group (**Table e1**).

HRCT Characteristics

Among all included subjects, 60.4% had excellent scan quality, 72.8% had a volumetric protocol, 82.3% had and expiratory series images, 87.4% had prone series images (**Table e2**). The maximum inspiratory series slice thickness included was 1.4mm.

Inter-rater agreement across 3 radiologists was good for dichotomized honeycombing, intermediate for mosaic attenuation, air trapping, traction bronchiectasis, axial, and craniocaudal distribution, and poor for reticulation, ground glass, and centrilobular nodules (**Table 2**). There was systematic variation by reader, with Reader 3 identifying more ground glass and less reticulation than reader 1 or 2, and Reader 1 identifying less centrilobular nodules and traction bronchiectasis than reader 2 or 3 (**Table e3**). **Table e4** shows weighted Kappa estimates for each pair of readers. **Table 3** and **Table e5** summarize the number and percentage of patients having dichotomized or categorical consensus features, and means and standard deviations for the 3-reader average of semi-quantitative scores, for the UMHS and LTRC cohorts, respectively.

Derivation of the Radiologic Diagnostic Model

In univariable analysis female sex, never smoking history, fewer pack-years smoked, absence of dichotomized honeycombing, absent or less extensive reticulation, absent or less extensive traction bronchiectasis and diffuse axial or craniocaudal distributions of abnormality were associated with increased odds of HP. Presence of or more extensive ground glass,

mosaic attenuation, air trapping, combined mosaic attenuation or air trapping (MA-AT), more extensive ground glass than reticulation (GG>Reticulation), more extensive mosaic attenuation or air trapping than reticulation (MA-AT>Reticulation), and centrilobular nodules were also associated with increased HP odds (**Table 4**). For simplicity, we limited candidate multivariable models to three categorical radiologic predictors. Several candidate models, adjusted for age, sex, and smoking status, were evaluated (**Table e6**). All had good discrimination (AUC>0.790). Because Model 7 (including radiologic MA-AT>reticulation and diffuse axial distribution) had a high AUC, maximized sensitivity for HP (sensitivity=65.3%) at the 90% specificity threshold, and had good face value, it was selected. Next, an HP-HRCT Diagnosis Score was created based on Model 7. In the adjusted model for the HP-HRCT Diagnosis Score (**Table 5** gives calculation parameters; score range 0-3), each 1-point score increase was associated with 2.45-fold increase in the odds of HP (CI 95% 1.99-3.02, p<0.0001; AUC=0.814). **Figure 3** shows an example HRCT for a UMHS HP patient with HP-HRCT Diagnosis Score of 3.

Model Validation and HP-HRCT Diagnosis Score Test Characteristics

As shown in **Table 6**, when mosaic attenuation or air trapping are more extensive than reticulation and disease has a diffuse axial distribution (i.e. when 3 of 3 points are assigned), there is a high specificity for HP (UMHS specificity 90.2% and LTRC specificity 95.8%). Sensitivity for HP when HP-HRCT Score = 3 was 55.4% in the UMHS Cohort and 18.2% in the LTRC Cohort. **Table e7** shows the final adjusted model in the LTRC cohort; the radiologic variables maintained predictive ability for HP in a similar strength and direction as the derivation model. **Figure e2** shows complete and HP-HRCT Diagnosis Score-based ROC curves for the UMHS and LTRC cohorts; AUC was > 0.70 in all models. The model was well-calibrated, with most probability-group points falling near the line of perfect calibration (**Figure e3**). The results of sensitivity analyses (**Table e8**) suggest that the HP-HRCT Diagnosis Score also identifies HP when it is defined by “classic” surgical biopsy features among all comparators and among comparators with idiopathic interstitial pneumonias, and identifies HP in the subgroup of

subjects without a documented exposure history. For reference, **Table e9** gives summary demographic and HRCT characteristics for the subset of HP patients (n=64) with a “classic” surgical lung biopsy.

Table e10 shows the regression formula allowing calculation of an adjusted model-based HP probability, and the positive predictive value (PPV) of various HP probability thresholds in the UMHS and LTRC cohorts. HP prevalence is higher in the UMHS than the LTRC cohort, thus, the same model-based HP probability corresponds to a higher PPV in the UMHS cohort. When model-based probability is 80% or greater, the PPV for HP is 100% in the UMHS and LTRC cohorts, and at a probability of 70% or greater UMHS PPV is 77% and LTRC PPV 46%. The PPV can be interpreted as the probability that a patient with a given test result (here, model-based HP probability) has HP. For comparison, an HP-HRCT Diagnosis Score of 3 corresponds to PPV=74% and NPV=80% in the UMHS cohort, and PPV=44% and NPV=86% in the LTRC cohort.

DISCUSSION

We describe the development and validation of a radiologic diagnosis model for HP. When a combination of diffuse axial distribution of interstitial abnormality and the extent of mosaic attenuation or air trapping is greater than that of reticulation, the risk of making a false positive HP diagnosis is less than 10%. Misdiagnosed UMHS patients (n=23) had desquamative interstitial pneumonia (n=3), IPF (n=10), Pneumoconiosis (n=1), respiratory bronchiolitis-interstitial lung disease (n=5), pulmonary vasculitis (n=2), non-specific interstitial pneumonitis (n=1), and sarcoidosis (n=1). Misdiagnosed LTRC patients (n=15) had IPF (n=4), non-specific interstitial pneumonitis (n=4), desquamative interstitial pneumonia (n=2), respiratory bronchiolitis-interstitial lung disease (n=2), uncharacterized fibrosis (n=1) and autoimmune disease (n=2).

Several previous analyses have evaluated the use of HRCT as a diagnostic test for HP in various populations. Lynch and colleagues retrospectively evaluated the CTs of 19 HP and 33

IPF patients and determined that an expert thoracic radiologist's CT diagnosis of definite HP was correct 92% of the time.(3) A "definite HP" pattern was not defined *a priori* in this study. HRCT features associated with HP included less honeycombing and traction bronchiectasis, and presence of micronodules and relative sparing of the lower half of the lower lung zone. Mosaic attenuation/perfusion was not evaluated. Silva and colleagues performed a similar analysis in a population of 16 HP, 23 IPF, and 25 NSIP patients.(4) Here, a confident CT diagnosis of HP was correct 88% of the time. HP patients had more lobular areas of decreased attenuation, air trapping, centrilobular nodules, relative sparing of the bases, random craniocaudal disease distribution, and peribronchovascular or random axial disease distribution. Johannson and colleagues evaluated HRCTs of a broad range of ILD diagnosis (HP prevalence 53%).(5) Clinical prediction models incorporating age, down feather exposure, bird exposure and either a radiologist's moderate to high confidence in an HP diagnosis, or HRCT features of diffuse craniocaudal distribution of ground glass abnormality and mosaic perfusion had good discriminative performance (c-statistic 0.758-0.778). Our model 4 in Table e6 is similar to the model of Johannson and colleagues, but with adjustment variables of sex and smoking history rather than bird/feather exposure. We found that ground glass as a dichotomous variable was not independently associated with HP, but mosaic attenuation (alone and when combined with air trapping or relative extent of reticulation) and diffuse axial and craniocaudal distributions are. Our HRCT evaluation methodology does not allow for exact replication of the model of Johannson and colleagues.

Our study expands upon the previously mentioned analyses in several ways. First, we include a broad range of non-HP ILD diagnoses, including fibrotic and non-fibrotic interstitial lung diseases. The HP-HRCT Diagnosis Score can therefore be applied across ILD phenotypes (i.e. fibrotic and non-fibrotic). Roughly half of our HP patients had fibrotic features of reticulation or traction bronchiectasis on HRCT. Second, the sensitivity analyses indicates that the HP-HRCT Diagnosis Score is also useful for identification of HP diagnosed when surgical lung

biopsy confirmed the presence of a classical/typical histopathologic HP pattern (as documented by the original interpreting University of Michigan pathologist). A challenging aspect of diagnosing a patient with ILD with HP is the idea that HP may present a variety of histopathologic patterns.(18-21) Our HP-HRCT Diagnosis Score therefore identifies a subset of ILD likely to have typical histopathologic findings of HP were a biopsy to be performed. Third, a sensitivity analysis excluding subjects with a documented exposure history finds the model to be robust in the group without identified environmental exposure. Fourth, we have validated our final model and a simple model-based HP-HRCT Diagnosis Score in a separate, multicenter cohort of ILD subjects. The UMHS cohort HRCTs were evaluated by a different methodology than those of the LTRC cohort. This supports the idea that our model-based score is useful as a simple dichotomous checklist that can be applied in practice even when a complex HRCT scoring methodology is not in use.

Application of this type of score in practice should involve consideration of the sensitivity and specificity of the test as presented here, as well as post-test probability of disease which is dependent on disease prevalence. With this in mind, several limitations of our HP-HRCT score should be noted. First, the sensitivity of a high HP-HRCT Diagnosis Score (>2 points) is relatively low (55.4% in the UMHS Cohort and 18.2% in the LTRC Cohort), indicating that a substantial fraction of patients with HP are not identified by a high HP-HRCT Diagnosis Score. Second, in applying this Score in practice the evaluating clinician should consider the probability of having HP given a positive test (i.e. the Positive Predictive Value [PPV] of an HP-HRCT Diagnosis Score of 3) when deciding whether additional testing is needed to make a confident HP diagnosis.(1) We present specificity of various HP-HRCT Diagnosis Score thresholds, which is not dependent on disease prevalence, and conclude that the risk of a false HP diagnosis is less than 10% when a patient receives all 3 points. However, PPV at the same threshold depends on disease prevalence which should be determined by clinicians using the test and based on knowledge of regional prevalence of HP or other characteristics of the patient such as

exposure history. In the UMHS cohort (HP prevalence 34%), an HP-HRCT Score of 3 is associated with PPV=74%. In the LTRC cohort (HP prevalence 16%), a Score of 3 has PPV=44%. It is unknown how additional clinical variables or diagnostic test results such as identification of an HP exposure or fluid lymphocytes during bronchoalveolar lavage will modify the probability/PPV of HP in concert with the HP-HRCT Diagnosis Score. These questions should be the subject of future, ideally prospective, studies. Decisions regarding the need for additional diagnostic testing such as bronchoscopy or surgical lung biopsy after finding a "positive" result using the HP-HRCT Diagnosis Score should be individualized to the patient and practice setting.

Interestingly, the most recent iteration of multi-society diagnosis guidelines for IPF includes "ground glass extent > reticular extent" as a feature inconsistent with radiologic UIP.(2) In our study, this variable was associated with HP (and therefore against IPF, given the makeup of the non-HP control population) in unadjusted analysis, but was not an independent predictor of HP after adjusting for age, sex, smoking history, presence of mosaic attenuation or air trapping, and axial disease distribution (**Table e6**, model 5). Finding more extensive mosaic attenuation or air trapping than reticulation is consistently and strongly associated with HP in our study.

Our study has several weaknesses. First, three thoracic radiologists scored each HRCT, generating robust data on agreement on HRCT features. This data alerted us to poor agreement on HRCT features of reticulation, ground glass, and centrilobular nodules. We did include reticulation in our model via the combined "MA-AT>Reticulation" variable, raising potential concerns about reproducibility of results if broadly applied. Despite poor agreement on presence of reticulation among UMHS radiologists, this reticulation-based variable remained strongly associated with an HP diagnosis in the LTRC cohort and the model appears to be valid despite measurement of reticulation using a different scoring system applied to CTs by different radiologists. Second, histopathologic disease confirmation was required for inclusion in the

control group of the UMHS cohort, and for HP and control groups of the LTRC cohort. This methodology may select for subjects with an atypical HRCT, thereby potentially altering conclusions about what features distinguish HP from other ILDs. The LTRC cohort included final diagnoses of unclassifiable fibrosis, mitigating this concern given good model performance in the validation cohort. Third, we can't rule out the possibility that HRCT findings had some influence on the final clinical diagnoses. While we were careful to blind clinical diagnosis verification to the HRCT reports/findings, these reports were available to the clinicians and multidisciplinary teams assigning the original clinical diagnoses to the UMHS and LTRC patients. Fourth, UMHS was a participating center in the LTRC study. Review of internal records indicates that 110 subjects from the UMHS cohort were enrolled in the LTRC study. The LTRC data was de-identified and did not include notation of the referring center, so we are unable to remove overlap. If all 110 UMHS cohort subjects were included among the 424 validation subjects, up to 26% of the validation cohort would overlap with the derivation cohort.

In conclusion, we have developed and validated a radiologic diagnosis model for HP such that a when the extent of mosaic attenuation or air trapping is greater than reticulation and diffuse axial distribution of interstitial abnormality are present in combination, the specificity for HP is greater than 90% (less than 10% false positive rate) for clinically-diagnosed hypersensitivity pneumonitis.

ACKNOWLEDGMENTS

This study utilized data provided by the Lung Tissue Research Consortium (LTRC) supported by the National Heart, Lung, and Blood Institute (NHLBI).

REFERENCES

1. Salisbury ML, Myers JL, Belloli EA, Kazerooni EA, Martinez FJ, Flaherty KR. Diagnosis and Treatment of Fibrotic Hypersensitivity Pneumonia. Where We Stand and Where We Need to Go. Am J Respir Crit Care Med. 2017;196(6):690-9.
2. Raghu G, Collard HR, Egan JJ, Martinez FJ, Behr J, Brown KK, 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.
3. Lynch DA, Newell JD, Logan PM, King TE, Jr., Muller NL. Can CT distinguish hypersensitivity pneumonitis from idiopathic pulmonary fibrosis? AJR Am J Roentgenol. 1995;165(4):807-11.
4. Silva CI, Muller NL, Lynch DA, Curran-Everett D, Brown KK, Lee KS, et al. Chronic hypersensitivity pneumonitis: differentiation from idiopathic pulmonary fibrosis and nonspecific interstitial pneumonia by using thin-section CT. Radiology. 2008;246(1):288-97.
5. Johannson KA, Elicker BM, Vittinghoff E, Assayag D, de Boer K, Golden JA, et al. A diagnostic model for chronic hypersensitivity pneumonitis. Thorax. 2016.
6. Welker L, Jorres RA, Costabel U, Magnussen H. Predictive value of BAL cell differentials in the diagnosis of interstitial lung diseases. Eur Respir J. 2004;24(6):1000-6.
7. Ohshimo S, Bonella F, Cui A, Beume M, Kohno N, Guzman J, et al. Significance of bronchoalveolar lavage for the diagnosis of idiopathic pulmonary fibrosis. Am J Respir Crit Care Med. 2009;179(11):1043-7.
8. Society AT. Idiopathic pulmonary fibrosis: diagnosis and treatment. International consensus statement. American Thoracic Society (ATS), and the European Respiratory Society (ERS). Am J Respir Crit Care Med. 2000;161(2 Pt 1):646-64.
9. Flaherty KR, King TE, Jr., Raghu G, Lynch JP, 3rd, Colby TV, Travis WD, et al. Idiopathic interstitial pneumonia: what is the effect of a multidisciplinary approach to diagnosis? Am J Respir Crit Care Med. 2004;170(8):904-10.
10. Salisbury ML, Xia M, Murray S, Bartholmai BJ, Kazerooni EA, Meldrum CA, et al. Predictors of idiopathic pulmonary fibrosis in absence of radiologic honeycombing: A cross sectional analysis in ILD patients undergoing lung tissue sampling. Respir Med. 2016;118:88-95.
11. Salisbury MG, BH; Chughtai, A; Sayyouh, M; Kazerooni, EA; Gu, T; Xia, M; Murray, S; Myers, JL; Lagstein, A; Konopka, KE; Belloli, EA; Sheth, JS; White, ES; Holtze, C; Flaherty, KR. . Utility of High-Resolution Computed Tomography for Diagnosis of Hypersensitivity Pneumonia. American Thoracic Society International Conference. Washington, D.C.; 2017.
12. Hansell DM, Bankier AA, MacMahon H, McLoud TC, Muller NL, Remy J. Fleischner Society: glossary of terms for thoracic imaging. Radiology. 2008;246(3):697-722.
13. Conger AJ. Integration and Generalization of Kappas for Multiple Raters. Psychological Bulletin. 1980;88(2):322-8.
14. Marasini D, Quatto P, Ripamonti E. Assessing the inter-rater agreement for ordinal data through weighted indexes. Stat Methods Med Res. 2014.

15. Mc NQ. Note on the sampling error of the difference between correlated proportions or percentages. Psychometrika. 1947;12(2):153-7.
16. Harrell FE, Jr., Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Stat Med. 1996;15(4):361-87.
17. Freeman EA, Moisen G. PresenceAbsence: An R package for presence absence analysis. Journal of Statistical Software. 2008;23(11):1-31.
18. Vourlekis JS, Schwarz MI, Cool CD, Tuder RM, King TE, Brown KK. Nonspecific interstitial pneumonitis as the sole histologic expression of hypersensitivity pneumonitis. Am J Med. 2002;112(6):490-3.
19. Ohtani Y, Saiki S, Kitaichi M, Usui Y, Inase N, Costabel U, et al. Chronic bird fancier's lung: histopathological and clinical correlation. An application of the 2002 ATS/ERS consensus classification of the idiopathic interstitial pneumonias. Thorax. 2005;60(8):665-71.
20. Churg A, Sin DD, Everett D, Brown K, Cool C. Pathologic patterns and survival in chronic hypersensitivity pneumonitis. Am J Surg Pathol. 2009;33(12):1765-70.
21. Chiba S, Tsuchiya K, Akashi T, Ishizuka M, Okamoto T, Furusawa H, et al. Chronic Hypersensitivity Pneumonitis With a Usual Interstitial Pneumonia-Like Pattern: Correlation Between Histopathologic and Clinical Findings. Chest. 2016;149(6):1473-81.

FIGURE LEGENDS

Figure 1. Derivation Cohort (University of Michigan/UMHS) Patient Flow Diagram

* Includes those with ILD diagnosis confirmed by HRCT only (i.e. “Definite UIP” HRCT pattern) and those unwilling or unable to undergo diagnostic biopsy

** Excluded diagnoses include primary cystic lung disease, PAP, airway or pulmonary vascular disease without ILD, drug induced lung disease, infection, and other non-ILD diagnoses (congestive heart failure, etc).

*** Of included patients identified from the ICD-9 search alone, 23 HP and 7 control patients had chart documentation of multidisciplinary case review which occurred prior to creation of a searchable database of conference minutes.

COP=Cryptogenic Organizing Pneumonia; RB-ILD=Respiratory Bronchiolitis-Interstitial Lung Disease; NSIP=Nonspecific Interstitial Pneumonia; CTD-ILD=Connective tissue disease associated ILD; DIP=Desquamative Interstitial Pneumonia; PLCH=Pulmonary Langerhans Cell Histiocytosis; HRCT=High Resolution Computed Tomography. HRCT=High Resolution Computed Tomography

Figure 2. Validation (LTRC) Cohort Patient Flow Diagram

COP=Cryptogenic Organizing Pneumonia; RB-ILD=Respiratory Bronchiolitis-Interstitial Lung Disease; NSIP=Nonspecific Interstitial Pneumonia; CTD-ILD=Connective tissue disease associated ILD; DIP=Desquamative Interstitial Pneumonia; PLCH=Pulmonary Langerhans Cell Histiocytosis; HRCT=High Resolution Computed Tomography

Figure 3. Representative Images from two patients with hypersensitivity pneumonitis and HP-HRCT Diagnosis Score of 3

Panels A, B, and C are inspiratory coronal reconstruction, inspiratory axial supine, and expiratory axial supine images, respectively, from a subject with fibrotic HP. Disease is diffusely

distributed in the axial direction (white arrows). An example of reticulation is shows in the box, and asterisks (*) mark regions of relatively decreased attenuation (mosaic attenuation) on the inspiratory image (panel B) and air trapping on the expiratory image (Panel C).
Panels D, E, and F are inspiratory coronal reconstruction, inspiratory axial supine, and expiratory axial supine images, respectively, from a subject with non-fibrotic HP. Disease is diffusely distributed in the axial direction (white arrows). Asterisks (*) mark regions of relatively decreased attenuation (mosaic attenuation) on the inspiratory image (panel E) and air trapping on the expiratory image (Panel F).

Figure 1. Derivation Cohort (University of Michigan/UMHS) Patient Flow Diagram

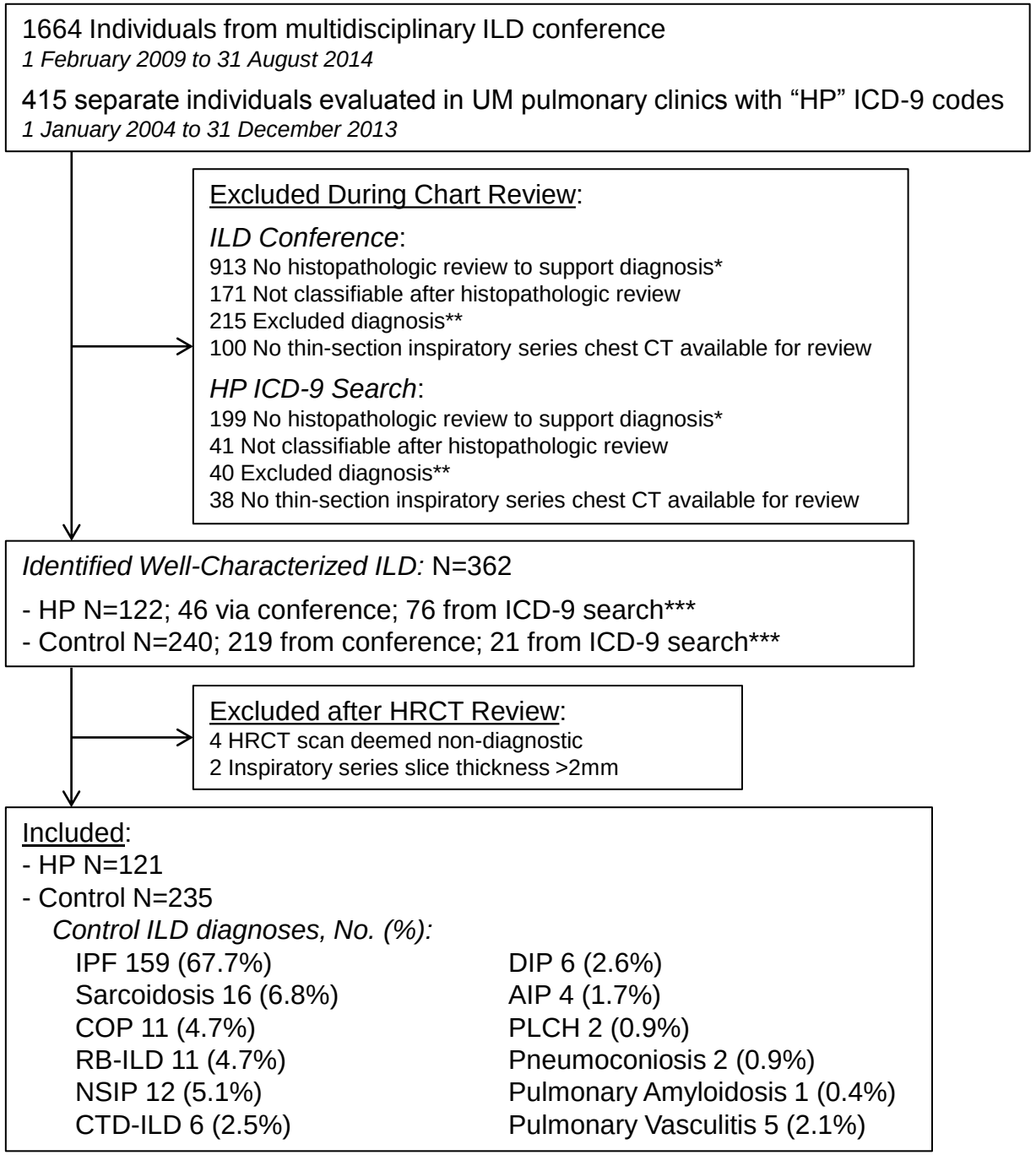


Figure 2. Validation (LTRC) Cohort Patient Flow Diagram

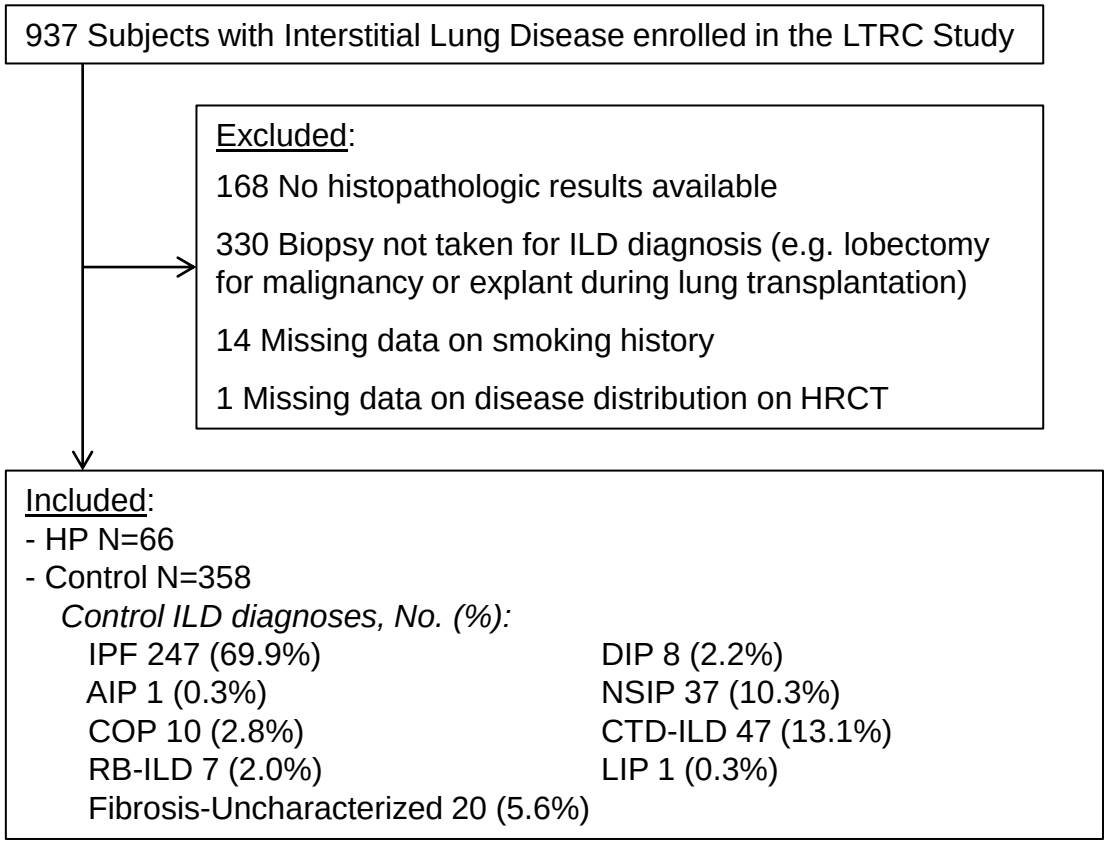
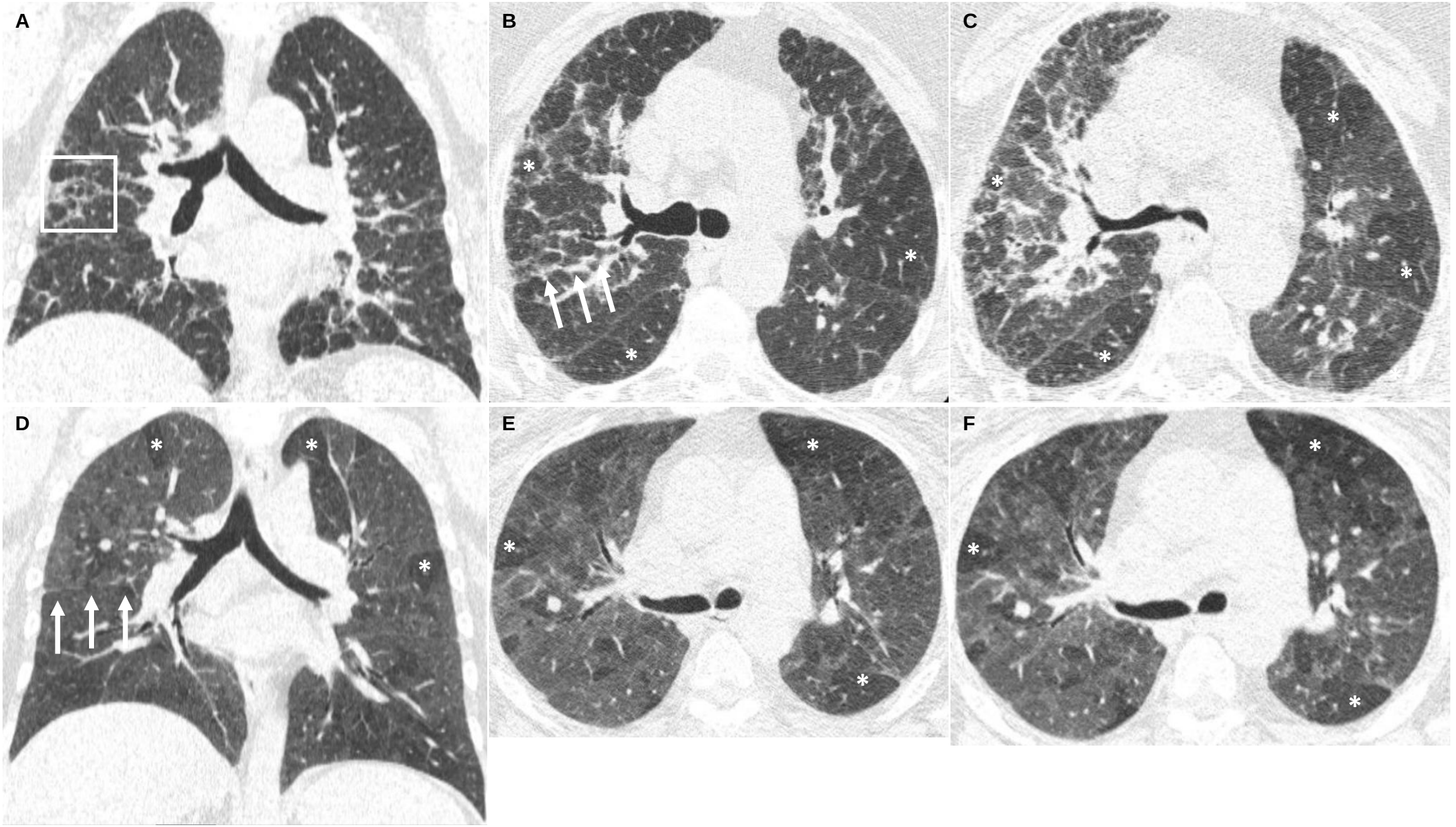


Figure 3.



TABLES

Table 1. Patient Characteristics at Baseline - UMHS Derivation Cohort

Variable	*All N=356*	*HP N=121*	*Not HP N=235*
Age, y, mean (SD)	59.6 (11.1)	58.1 (10.9)	60.3 (11.1)
Sex, No. (%) male	175 (49.2%)	39 (32.2%)	136 (57.9%)
Race, No. (%) white	308 (86.5%)	108 (89.3%)	200 (85.1%)
Smoking Status, No. (%)			
Current	30 (8.4%)	4 (3.3%)	26 (11.1%)
Former	167 (46.9%)	54 (44.6%)	113 (48.1%)
Never	157 (44.7%)	63 (52.1%)	96 (40.9%)
Pack-years Smoked, y, mean (SD) (n=5 missing)	14.6 (19.9)	11.0 (17.6)	16.4 (20.8)
Baseline Physiology, mean (SD)			
FVC %-Predicted (n=345)	66.0 (17.6)	65.7 (17.4)	66.1 (17.7)
FEV1 %-Predicted (n=345)	75.2 (19.9)	73.6 (20.5)	76.0 (19.5)
FEV1/FVC Ratio (n=345)	82.5 (8.0)	81.9 (8.4)	82.7 (7.7)
DLCO %-Pred (n=248)	52.0 (18.7)	51.4 (16.6)	52.3 (19.7)
Median (IQR) time between HRCT and PFT, days	8 (0-50)	7 (0-40)	11 (0-60)
Reported Symptom Duration, months, mean (SD)	25.3 (33.2)	27.5 (43.9)	24.2 (26.0)
Corticosteroid Prior to University of Michigan Evaluation, No. (%)	113 (32.0%)	46 (38.0%)	68 (28.9%)
Diagnostic Testing Available, No. (%)			
Surgical Lung Biopsy Performed	304 (85.4%)	86 (71.1%)	218 (92.8%)
"Classic HP" on surgical lung biopsy (among HP)	--	64 (52.9%)	--
Transbronchial Lung Biopsy Performed	111 (31.2%)	63 (52.1%)	48 (20.4%)
BAL Cell Count and Differential, No. (%) of total with test	91 (25.5%)	54 (44.6%)	37 (15.7%)
% Lymphs, mean (SD)	22.6 (20.4)	32.2 (19.5)	8.5 (11.5)
"HP Panel" Results, No. (%) of total with test	91 (25.6%)	66 (54.5%)	25 (10.6%)
Negative, No. (%) of those with test	59 (64.8%)	41 (62.1%)	18 (72.0%)
Bird, No. (%) of those with test	14 (15.4%)	12 (18.2%)	2 (8.0%)
Microbe, No. (%) of those with test	11 (12.1%)	7 (10.6%)	4 (16.0%)
Both, No. (%) of those with test	7 (7.7%)	6 (9.1%)	1 (4.0)
Reported Exposures No. (%)			
None reported	217 (61.0%)	38 (31.4%)	179 (76.2%)
Asbestos	17 (4.8%)	0 (0%)	17 (7.2%)
Beryllium	2 (0.6%)	0 (0%)	2 (0.9%)
Bird	35 (9.8%)	34 (28.1%)	1 (0.4%)
Chemical	5 (1.4%)	1 (0.8%)	4 (1.7%)
Drug	1 (0.3%)	1 (0.8%)	0 (0%)
Hot tub	10 (2.8%)	9 (7.4%)	1 (0.4%)
Industrial dust	14 (3.9%)	0 (0%)	14 (6.0%)
Microbe	35 (9.8%)	25 (20.7%)	10 (4.3%)
Multiple	11 (3.1%)	10 (8.3%)	1 (0.4%)
Wood dust	9 (2.5%)	3 (2.5%)	6 (2.6%)

Table 2. Dichotomous* 3-radiologist agreement – UMHS Cohort

HRCT Feature	2-way reader Kappa (95% CI)			Light's Kappa (95% CI)
	Reader 1 vs 2	Reader 1 vs 3	Reader 2 vs 3	
Honeycombing	0.61 (0.47-0.75)	0.59 (0.46-0.72)	0.63 (0.51-0.75)	0.61 (0.50-0.72)
Reticulation	0.53 (0.44-0.62)	0.26 (0.19-0.33)	0.14 (0.09-0.18)	0.31 (0.26-0.36)
Ground Glass	0.49 (0.40-0.59)	0.18 (0.09-0.27)	0.20 (0.12-0.29)	0.29 (0.23-0.36)
Mosaic Attenuation	0.59 (0.49-0.69)	0.66 (0.57-0.75)	0.53 (0.43-0.64)	0.59 (0.52-0.67)
Air Trapping (n=293 expiratory)	0.50 (0.41-0.59)	0.64 (0.56-0.73)	0.48 (0.38-0.58)	0.54 (0.47-0.61)
Centrilobular Nodules	0.18 (0.09-0.27)	0.47 (0.30-0.64)	0.34 (0.23-0.45)	0.33 (0.24-0.42)
Traction Bronchiectasis	0.50 (0.42-0.59)	0.59 (0.51-0.67)	0.66 (0.57-0.75)	0.59 (0.52-0.65)
Craniocaudal Distribution	0.55 (0.48-0.63)	0.50 (0.43-0.57)	0.47 (0.40-0.54)	0.51 (0.46-0.56)
Axial Distribution	0.45 (0.37-0.52)	0.47 (0.40-0.54)	0.34 (0.27-0.40)	0.42 (0.36-0.47)

*A feature was considered absent when the average score (for honeycombing, reticulation, ground glass, mosaic attenuation, air trapping) or sum score (centrilobular nodules, traction bronchiectasis) was less than 0.5, otherwise the feature was considered present.

Table 3. Consensus HRCT features – UMHS Cohort

Variable	Group		
	All Subjects (n=356)	HP (n=121)	Not HP (n=235)
Dichotomous Scores**, No. (%)			
Honeycombing	37 (10.4)	10 (8.3)	27 (11.5)
Reticulation	230 (64.6)	52 (43.0)	178 (75.7)
Ground Glass	256 (74.4)	103 (85.1)	162 (68.9)
Mosaic Attenuation	89 (25.0)	62 (51.2)	27 (11.5)
Air Trapping (n=293 with expiratory)	127 (43.3)	75 (71.4)	52 (27.7)
Mosaic Attenuation or Air Trapping (MA-AT)	145 (40.7)	85 (70.3)	60 (25.5)
GG > Reticulation	229 (64.3)	99 (81.8)	130 (55.3)
MA-AT > Reticulation	130 (36.5)	85 (70.3)	45 (19.2)
Centrilobular Nodules	31 (8.7)	16 (13.2)	15 (6.4)
Traction Bronchiectasis	252 (70.8)	60 (49.6)	192 (81.7)
Semi-quantitative (SQ) Scores*, Mean (SD)			
Honeycombing	0.20 (0.47)	0.16 (0.44)	0.22 (0.49)
Reticulation	0.86 (0.53)	0.63 (0.54)	0.98 (0.48)
Ground Glass	1.59 (1.06)	2.01 (1.22)	1.37 (0.90)
Mosaic Attenuation	0.50 (0.83)	1.04 (1.06)	0.22 (0.48)
Air Trapping (of n=293 w/ expiratory)	0.82 (0.87)	1.41 (0.93)	0.49 (0.63)
Centrilobular Nodules	0.46 (0.98)	0.71 (1.15)	0.34 (0.87)
Traction Bronchiectasis	2.84 (1.83)	1.92 (1.81)	3.31 (1.65)
Craniocaudal Distribution, probability (SD)			
Upper Lung	0.14 (0.28)	0.18 (0.30)	0.12 (0.27)
Lower Lung	0.55 (0.43)	0.31 (0.38)	0.67 (0.40)
Diffuse	0.31 (0.36)	0.50 (0.37)	0.22 (0.30)
Axial Distribution, probability (SD)			
Central Lung	0.08 (0.19)	0.07 (0.15)	0.08 (0.20)
Peripheral/Subpleural Lung	0.52 (0.42)	0.30 (0.38)	0.63 (0.40)
Peribronchovascular	0.01 (0.06)	0.003 (0.03)	0.02 (0.07)
Diffuse	0.39 (0.38)	0.63 (0.39)	0.27 (0.32)

*Semi-quantitative scores: For honeycombing, reticulation, ground glass, mosaic attenuation, air trapping, the average score was determined by summing 5 lobe scores and dividing by 5 for each subject and each radiologist/reader. For centrilobular nodules, traction bronchiectasis, the sum was determined by adding the score for 5 lobes. Average and sum scores are scaled 0-5, corresponding to proportion of total lung having the feature (for honeycombing, reticulation, ground glass, mosaic attenuation, air trapping) or number of lobes having the feature (centrilobular nodules, traction bronchiectasis). The 3 radiologist's scores were then averaged.
**Dichotomous scores: A feature was designated as absent for each radiologist/reader when the average score (for honeycombing, reticulation, ground glass, mosaic attenuation, air trapping) or sum score (centrilobular nodules, traction bronchiectasis) for that features was less than 0.5, otherwise the feature was considered present. The consensus scores are shown in this table. A feature was considered "present" when 2 of 3 radiologists called it present.
Mosaic attenuation or air trapping ("MA-AT") was "present" when a subject had consensus presence of mosaic attenuation or air trapping. The "GG>Reticulation" variable was present when the semi-quantitative ground glass score was higher than the semi-quantitative reticulation score. The "MA-AT>Reticulation" variable was present when the semi-quantitative score for mosaic attenuation or air trapping was higher than the semi-quantitative reticulation score.

Table 4. Univariable Logistic Regression Predicting HP Diagnosis – UMHS Cohort

	All Subjects	
Predictor	*OR (95% CI)*	*P*
Age, per 1 year	0.98 (0.96-1.002)	0.08
Male Sex	0.35 (0.22-0.55)	<0.001
Race (white vs others)	1.45 (0.74-2.86)	0.28
Ever Smoker Status (v never)	0.64 (0.41-0.99)	0.04
Pack-years Smoked, per +1 y	0.99 (0.97-0.997)	0.02
Symptom Duration, per +1 mo	1.00 (0.996-1.01)	0.38
Corticosteroid Prior to U-M	1.51 (0.95-2.39)	0.08
Baseline Physiology, per +10%		
FVC %-Predicted	0.99 (0.87-1.20)	0.83
FEV1 %-Predicted	0.94 (0.84-1.06)	0.30
FEV1/FVC Ratio	0.86 (0.65-1.13)	0.28
DLCO %-Pred	0.97 (0.85-1.12)	0.71
Honeycombing Present	0.47 (0.25-0.89)	0.02
Honeycombing Semi-Quantitative (SQ)*	0.73 (0.44-1.23)	0.24
Reticulation Present	0.24 (0.15-0.39)	<0.001
Reticulation SQ*	0.26 (0.16-0.41)	<0.001
Ground Glass Present	2.58 (1.46-4.57)	0.001
Ground Glass SQ*	1.78 (1.43-2.22)	<0.001
Mosaic Attenuation Present	8.10 (4.73-13.84)	<0.001
Mosaic Attenuation SQ*	3.88 (2.71-5.55)	<0.001
Air Trapping Present	6.54 (3.85-11.11)	<0.001
Air Trapping SQ*	4.06 (2.85-5.79)	<0.001
MA-AT Present	6.89 (4.23-11.22)	<0.001
GG > Reticulation Present	3.64 (2.42-6.17)	<0.001
MA-AT > Reticulation Present	9.97 (6.00-16.56)	<0.001
Centrilobular Nodules Present	2.24 (1.06-4.69)	0.03
Centrilobular Nodules SQ*	1.45 (1.16-1.81)	0.001
Traction Bronchiectasis Present	0.22 (0.14-0.36)	<0.001
Traction Bronchiectasis SQ*	0.65 (0.57-0.74)	<0.001
Craniocaudal Distribution		
Upper Lung	0.31 (0.12-0.78)	0.01
Lower Lung	0.08 (0.04-0.16)	<0.001
Diffuse	*Ref*	*Ref*
Axial Distribution		
Central	0.14 (0.03-0.61)	0.01
Peripheral/Subpleural	0.08 (0.04-0.15)	<0.001
Peribronchovascular	0.0007 (0-0.38)	0.02
Diffuse	*Ref*	*Ref*

*ORs for semi-quantitative scores are per 1-unit increase in the score. Scores are on a scale of 0 to 5.

SQ=Semi-quantitative. Mosaic attenuation or air trapping ("MA-AT") was "present" when a subject had consensus presence of mosaic attenuation or air trapping. The "GG>Reticulation" variable was present when the semi-quantitative ground glass score was higher than the semi-quantitative reticulation score. The "MA-AT>Reticulation" variable was present when the semi-quantitative score for mosaic attenuation or air trapping was higher than the semi-quantitative reticulation score.

Table 5. HP-HRCT Diagnosis Score Calculation

HRCT Feature	*Points*
MA or AT more extensive than Reticulation	2
Diffuse Axial Disease Distribution	1

MA: Mosaic Attenuation; AT: Air Trapping

Table 6. HP-HRCT Diagnosis Score Test Characteristics in UMHS and LTRC Cohorts

UMHS Cohort				
Threshold	*HP (n)*	*Not HP (n)*	*Sens (CI 95%)*	*Spec (CI 95%)*
>0	95	77	78.5 (71.2 - 85.8)	67.2 (61.2 - 73.2)
>1	85	45	70.2 (62.1 - 78.4)	80.9 (75.8 - 85.9)
>2	67	23	55.4 (46.5 - 64.2)	90.2 (86.4 - 94.0)
LTRC Cohort				
Threshold	*HP (n)*	*Not HP (n)*	*Sens (CI 95%)*	*Spec (CI 95%)*
>0	35	70	53.0 (41.0 - 65.1)	80.4 (76.3 - 84.6)
>1	22	29	33.3 (22.0 - 44.7)	91.9 (89.1 - 94.7)
>2	12	15	18.2 (8.9 - 27.5)	95.8 (93.7 - 97.9)

Figure e1. UMHS Standardized HRCT Score Sheet

HRCT SCORE SHEET

HRCT Protocol - Check what applies				
	Inspiratory	Expiratory	Prone	Quality (circle)
Slice Thickness (mm)		Y N	Y N	1 - Excellent 2 - Diagnostic 3 - Nondiagnostic
Volumetric?	Y N			Resp Motion
Inter-slice Interval (mm) **(if incremental)**				Patient Noise Other

Name #REF!

MRN #REF!

HRCT Date #REF!

Reader **BG** **MS** **AC**

Traction Bronchiectasis			**Centrilobular Nodules**		
Lobes	Absent; Present		Lobes	Absent; Present	
RUL	0	1	RUL	0	1
RML	0	1	RML	0	1
RLL	0	1	RLL	0	1
LUL/Lingula	0	1	LUL/Lingula	0	1
LLL	0	1	LLL	0	1

Craniocaudal Distribution		**Axial/Other Distribution**	
upper lungs		central	
lower lungs		peripheral	
diffuse		subpleural	
		peribronchovascular	
		diffuse	

SCALE		**SEMI-QUANTITATIVE SCORES**					
none	0	Lobes	Normal	Honeycombing	Reticulation	Ground Glass	Mosaic Attenuation	Air Trapping
<5% (minimal)	1	RUL	Y	0 1 2 3 4 5	0 1 2 3 4 5	0 1 2 3 4 5	0 1 2 3 4 5	0 1 2 3 4 5
<25% of lobe	2	RML	Y	0 1 2 3 4 5	0 1 2 3 4 5	0 1 2 3 4 5	0 1 2 3 4 5	0 1 2 3 4 5
25-49% of lobe	3	RLL	Y	0 1 2 3 4 5	0 1 2 3 4 5	0 1 2 3 4 5	0 1 2 3 4 5	0 1 2 3 4 5
50-75% of lobe	4	LUL	Y	0 1 2 3 4 5	0 1 2 3 4 5	0 1 2 3 4 5	0 1 2 3 4 5	0 1 2 3 4 5
>75% of lobe	5	LLL	Y	0 1 2 3 4 5	0 1 2 3 4 5	0 1 2 3 4 5	0 1 2 3 4 5	0 1 2 3 4 5

Findings	**Major**	**Minor**
interlobularseptal thickening		
intralobular **lines**		
subpleural **lines**		
parenchymal bands		
traction **bronchiectasis**		
non-traction **bronchiectasis**		
peribronchial thickening		
lobar or focal consolidation		
cysts		
miliary **nodules**		
centrilobular **nodules**		
peribronchovascular **nodules**		
subpleural **nodules**		
cavitary **nodules**		
tree-in-bud changes		
emphysema		
upper lobe **volume loss**		
lower lobe **volume loss**		
other _____		
other _____		

Diagnosis	**Choice**		**Confidence**			**Comments**
	1st	2nd	Def	Prob	Poss	
UIP						
NSIP						
Smoking Fibrosis						
Acute HP						
Chronic HP						
RB						
DIP						
COP						
Lymphangitic Spread						
LIP						
LAM						
PLCH (EG)						
Sarcoid						
Other (specify)						
Normal						

Miscellaneous
PA Diameter (mm) _____
Ao Diameter (mm) _____
Pleural Plaques: Y N
Lymphadenopathy: Y N
Subpleural Sparing: Y N

NOTES/COMMENTS

Table e1. Patient Characteristics at Baseline - LTRC Validation Cohort

Variable	*All N=438*	*HP N=67*	*Not HP N=371*
Age, y, mean (SD)	62.3 (9.7)	60.8 (10.7)	62.6 (9.5)
Sex, No. (%) male	229 (54.0)	24 (36.4)	205 (57.3)
Smoking Status, No. (%)			
Current	10 (2.4)	0 (0)	10 (2.8)
Former	247 (58.3)	38 (57.6)	209 (58.4)
Never	167 (39.4)	28 (42.4)	139 (38.8)
Pack-years Smoked (among ever smokers), y, mean (SD)	27.2 (23.2)	23.7 (21.4)	27.8 (23.5)
Baseline Physiology, mean (SD)			
FVC %-Predicted (n=415)	68.6 (15.8)	70.9 (15.2)	68.2 (15.9)
FEV1 %-Predicted (n=415)	73.7 (17.1)	73.1 (16.5)	73.8 (17.3)
DLCO %-Predicted (n=380)	57.1 (19.2)	64.5 (20.6)	55.7 (18.7)

Table e2. HRCT Quality Scores – UMHS Cohort

Variable	All N=356	HP N=121	Not HP N=235
Scan Quality, No. (%)			
Excellent	215 (60.4)	75 (62.0)	140 (59.6)
Diagnostic*	141 (39.6)	46 (38.0)	95 (40.4)
Inspiratory Slice Thickness, mm, Median (range)	1.25 (0.63, 1.40)	1.25 (0.63, 1.25)	1.25 (0.63, 1.40)
Volumetric Protocol Available, No. (%)	259 (72.8)	81 (66.9)	178 (75.7)
Inter-slice Interval** in mm, Median (range)	10 (1, 20)	10 (1, 20)	10 (1.25, 20)
Expiratory Series Available, No. (%)	293 (82.3)	105 (86.8)	188 (80.0)
Prone Series Available, No. (%)	311 (87.4)	104 (86.0)	207 (88.1)

*Scan quality was not “excellent” but was sufficiently good to render a radiologic diagnostic opinion.
**For those HRCTs without a volumetric protocol

Table e3. Semi-quantitative and dichotomous scores averaged over all HRCTs, shown by radiologist – UMHS Cohort

Variable	Reader		
	1	2	3
Semi-quantitative Scores, Mean (SD)			
Honeycombing	0.12 (0.35); Pvs2 < 0.001	0.18 (0.40); Pvs3 < 0.001	0.31 (0.79); Pvs1 < 0.001
Reticulation	0.76 (0.59); Pvs2 < 0.001	1.52 (0.93); Pvs3 < 0.001	0.29 (0.49); Pvs1 < 0.001
Ground Glass	1.25 (1.012); Pvs2 = 0.02	1.16 (1.10); Pvs3 < 0.001	2.33 (1.35); Pvs1 < 0.001
Mosaic Attenuation	0.49 (0.86); Pvs2 = 0.006	0.39 (0.70); Pvs3 < 0.001	0.63 (1.22); Pvs1 < 0.001
Air Trapping (of n=293 expiratory)	0.95 (0.85); Pvs2 < 0.001	0.80 (0.77); Pvs3 < 0.001	1.07 (1.22); Pvs1 < 0.001
Centrilobular Nodules	0.14 (0.72); Pvs2 < 0.001	0.82 (1.56); Pvs3 < 0.001	0.44 (1.35); Pvs1 < 0.001
Traction Bronchiectasis	1.96 (2.03); Pvs2 < 0.001	3.33 (2.05); Pvs3 = 0.23	3.23 (2.23); Pvs1 < 0.001
Dichotomous Scores, No. (%)			
Honeycombing	39 (9%); Pvs2 = 0.16	39 (11%); Pvs3 = 0.01	53 (15%); Pvs1 = <0.001
Reticulation	233 (66%); Pvs2 < 0.001	287 (81%); Pvs3 < 0.001	96 (27%); Pvs1 < 0.001
Ground Glass	237 (67%); Pvs2 = 0.10	222 (62%); Pvs3 < 0.001	320 (90%); Pvs1 < 0.001
Mosaic Attenuation	98 (28%); Pvs2 = 0.11	86 (24%); Pvs3 = 0.90	87 (24%); Pvs1 = 0.11
Air Trapping (of n=293 expiratory)	102 (35%); Pvs2 < 0.001	159 (54%); Pvs3 = 0.09	144 (49%); Pvs1 < 0.001
Centrilobular Nodules	15 (4%); Pvs2 < 0.001	95 (27%); Pvs3 < 0.001	37 (10%); Pvs1 < 0.001
Traction Bronchiectasis	201 (56%); Pvs2 < 0.001	277 (78%); Pvs3 = 0.005	258 (72%); Pvs1 < 0.001
Craniocaudal Distribution, No. (%)****	P vs 2 = 0.85	P vs 3 < 0.001	P vs 1 < 0.001
Upper Lung	34 (9)	40 (11)	73 (20)
Lower Lung	180 (51)	183 (51)	220 (62)
Diffuse	141 (39)	132 (37)	63 (18)
Axial Distribution, No. (%)****	P vs 2 < 0.001	P vs 3 < 0.001	P vs 1 < 0.001
Central Lung	14 (4)	57 (16)	11 (3)
Peripheral Lung	183 (52)	136 (38)	211 (59)
Subpleural	0 (0)	1 (0)	22 (6)
Peribronchovascular	0 (0)	3 (1)	11 (3)
Diffuse	158 (44)	148 (44)	101 (28)

*Semi-quantitative scores: For honeycombing, reticulation, ground glass, mosaic attenuation, air trapping, the average score was determined by summing 5 lobe scores and dividing by 5. For centrilobular nodules, traction bronchiectasis, the sum was determined by adding the score for 5 lobes. Average and sum scores are scaled 0-5, corresponding to proportion of total lung having the feature (for honeycombing, reticulation, ground glass, mosaic attenuation, air trapping) or number of lobes having the feature (centrilobular nodules, traction bronchiectasis).

**Dichotomous scores: A feature was considered absent when the average score (for honeycombing, reticulation, ground glass, mosaic attenuation, air trapping) or sum score (centrilobular nodules, traction bronchiectasis) was less than 0.5, otherwise the feature was considered present.

***P-value gives a paired t-test comparing the mean score (semi-quantitative) or the McNemar's Chi-squared test for non-independent samples (dichotomous) for the indicated pair of readers

****Reader 1 and 2 deemed one scan each to be normal and did not score disease distribution.

Table e4. Semi-quantitative* 3-reader agreement (Weighted Kappa) – UMHS Cohort

Variable	2-way reader Weighted Kappa (95% CI)		
	1 vs 2	1 vs 3	2 vs 3
Honeycombing	0.71 (0.57-0.86)	0.57 (0.42-0.72)	0.52 (0.37-0.66)
Reticulation	0.35 (0.28-0.42)	0.06 (0.02-0.10)	0.17 (0.10-0.25)
Ground Glass	0.70 (0.63-0.76)	0.42 (0.34-0.50)	0.43 (0.35-0.51)
Mosaic Attenuation	0.63 (0.53-0.73)	0.55 (0.46-0.65)	0.71 (0.64-0.78)
Air Trapping (n=293 expiratory)	0.65 (0.57-0.73)	0.58 (0.51-0.66)	0.69 (0.61-0.76)
Centrilobular Nodules	0.15 (0.04-0.25)	0.32 (0.19-0.45)	0.48 (0.26-0.69)
Traction Bronchiectasis	0.50 (0.39-0.60)	0.72 (0.61-0.83)	0.59 (0.49-0.69)

*For honeycombing, reticulation, ground glass, mosaic attenuation, air trapping, the average score was determined by summing 5 lobe scores and dividing by 5, then rounding to the nearest integer. For centrilobular nodules, traction bronchiectasis, the sum was determined by adding the score for 5 lobes. Each are scaled 0-5 corresponding to proportion of total lung having the feature (for honeycombing, reticulation, ground glass, mosaic attenuation, air trapping) or number of lobes having the feature (centrilobular nodules, traction bronchiectasis).

Table e5. Summary of HRCT features – LTRC Cohort

Variable	Group		
	All Subjects (n=424)	HP (n=66)	Not HP (n=358)
Dichotomous Scores, No. (%)			
Honeycombing	36 (8)	0 (0)	36 (10)
Reticulation	294 (69)	31 (47)	263 (73)
Ground Glass	179 (42)	34 (52)	145 (41)
Mosaic Attenuation	31 (7)	12 (18)	19 (5)
Air Trapping	45 (11)	18 (27)	27 (0.8)
Mosaic Attenuation or Air Trapping	65 (15)	22 (33)	43 (12)
GG>Reticulation	131 (30)	33 (52)	97 (27)
MA-AT>Reticulation	51 (12)	22 (33)	29 (8)
Centrilobular Nodules	17 (4)	5 (8)	12 (3)
Traction Bronchiectasis	77 (18)	6 (9)	71 (20)
Semi-quantitative Scores, Mean (SD)			
Honeycombing	0.09 (0.2)	0.01 (0.05)	0.10 (0.22)
Reticulation	0.77 (0.59)	0.45 (0.54)	0.83 (0.58)
Ground Glass	0.58 (0.78)	0.86 (0.99)	0.53 (0.72)
Mosaic Attenuation	0.11 (0.41)	0.27 (0.60)	0.08 (0.36)
Air Trapping	0.13 (0.38)	0.30 (0.54)	0.10 (0.34)
Centrilobular Nodules	0.06 (0.29)	0.10 (0.25)	0.05 (0.30)
Traction Bronchiectasis	0.21 (0.28)	0.11 (0.16)	0.23 (0.30)
Craniocaudal Distribution, No. (%)			
None	25 (6)	10 (15)	15 (4)
Upper Lung	29 (7)	9 (14)	20 (6)
Lower Lung	122 (52)	27 (41)	195 (54)
Diffuse	148 (35)	20 (3)	128 (36)
Axial Distribution, No. (%)			
None	26 (6)	10 (15)	16 (4)
Peripheral/Subpleural	312 (74)	30 (45)	282 (79)
Central Lung	5 (1)	1 (2)	4 (1)
Even/Diffuse	81 (19)	25 (38)	56 (16)

Table e6. Candidate Multivariable Models – UMHS Cohort

	Model 1 AUC = 0.792 Spec=0.90 / Sens=0.529		Model 2 AUC = 0.803 Spec=0.90 / Sens=0.545		Model 3 AUC = 0.800 Spec=0.90 / Sens=0.562		Model 4 AUC = 0.795 Spec=0.90 / Sens=0.496	
Variable	OR (CI 95%)	P	OR (CI 95%)	P	OR (CI 95%)	P	OR (CI 95%)	P
Age	0.99 (0.96-1.02)	0.58	0.99 (0.97-1.02)	0.60	1.00 (0.98-1.03)	0.94	0.98 (0.96-1.003)	0.09
Male	0.51 (0.30-0.85)	0.01	0.52 (0.31-0.87)	0.01	0.51 (0.30-0.87)	0.01	0.50 (0.30-0.83)	0.01
Ever Smoker	0.72 (0.43-1.20)	0.21	0.69 (0.41-1.16)	0.17	0.73 (0.43-1.22)	0.23	0.70 (0.42-1.16)	0.17
Ground glass	1.85 (0.94-3.64)	0.08	--	--	--	--	1.33 (0.70-2.54)	0.39
MA or AT	4.43 (2.56-7.67)	<0.001	3.74 (2.11-6.62)	<0.001	3.51 (1.94-6.35)	<0.001	4.24 (2.41-7.47)	<0.001
GG>Reticulation	--	--	--	--	--	--	--	--
MA-AT>Reticulation	--	--	--	--	--	--	--	--
Traction Bronchiectasis	0.33 (0.18-0.61)	<0.001	0.41 (0.23-0.73)	0.003	0.39 (0.22-0.71)	0.002	--	--
Craniocaudal Diffuse	--	--	2.32 (1.30-4.12)	0.004	--	--	2.45 (1.40-4.35)	0.002
Axial Diffuse	--	--	--	--	2.28 (1.26-4.13)	0.007	--	--
	Model 5 AUC = 0.794 Spec=0.90 / Sens=0.620		Model 6 AUC = 0.818 Spec=0.90 / Sens=0.653		**Model 7** **AUC = 0.814** **Spec=0.90 / Sens=0.653**		Model 8 AUC = 0.791 Spec=0.90 / Sens=0.488	
Age	0.99 (0.97-1.01)	0.38	0.99 (0.96-1.01)	0.27	**0.99 (0.97-1.02)**	**0.60**	0.99 (0.97-1.01)	0.30
Male	0.50 (0.30-0.83)	0.008	0.49 (0.29-0.84)	0.009	**0.50 (0.29-0.85)**	**0.01**	0.51 (0.30-0.86)	0.008
Ever Smoker	0.75 (0.45-1.26)	0.28	0.78 (0.46-1.32)	0.35	**0.81 (0.48-1.37)**	**0.43**	0.73 (0.44-1.22)	0.22
Ground glass	--	--	--	--	**--**	**--**	1.16 (0.60-2.25)	0.50
MA or AT	3.89 (2.15-7.04)	<0.001	--	--	**--**	**--**	3.45 (1.89-6.30)	<0.001
GG>Reticulation	1.47 (0.79-2.76)	0.23	--	--	**--**	**--**	--	--
MA-AT>Reticulation	--	--	6.39 (3.67-11.13)	<0.001	**6.20 (3.53-10.90)**	**<0.001**	--	--
Traction Bronchiectasis	--	--	--	--	**--**	**--**	--	--
Craniocaudal Diffuse	--	--	2.36 (1.33-4.17)	0.003	**--**	**--**	--	--
Axial Diffuse	2.24 (1.23-4.07)	0.008	--	--	**2.33 (1.31-4.16)**	**0.004**	4.37 (1.94-9.83)	0.005

Footnotes: Due to collinearity, axial and craniocaudal distribution (Pearson corr=0.7 for the consensus probabilities) were not modeled together.

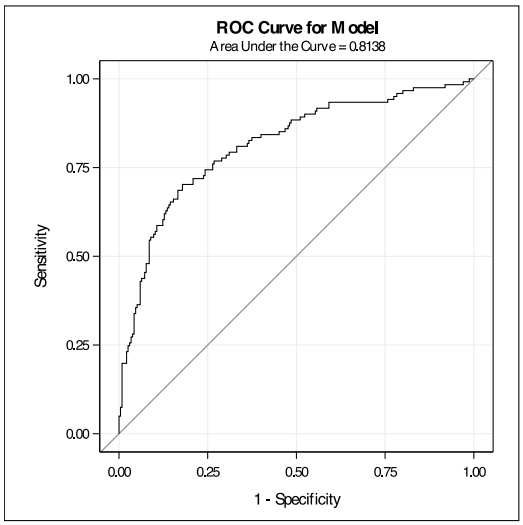
Table e7. Final Multivariable Model – LTRC Cohort

	Model 3 AUC = 0.708 Spec=0.90 / Sens=0.318	
Variable	OR (CI 95%)	P
Age	0.999 (0.97-1.03)	0.95
Male	0.60 (0.33-1.09)	0.10
Ever Smoker	1.08 (0.61-1.94)	0.78
MA-AT > Reticulation	3.66 (1.80-7.44)	<0.001
Axial Diffuse	2.11 (1.12-1.98)	0.02

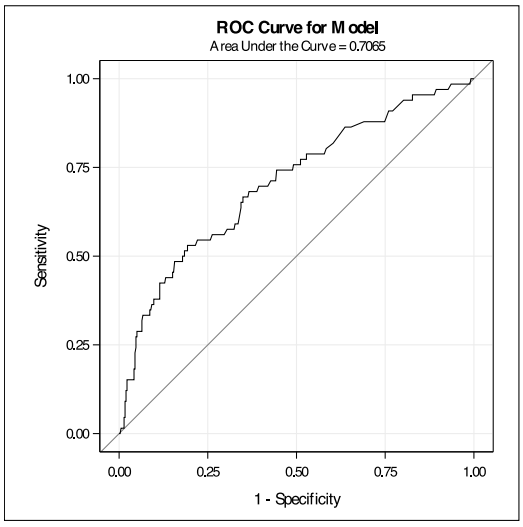
Figure e2. ROC Curves for Final Full and HP-HRCT Diagnosis Score-Based Models in UMHS and LTRC Cohorts

Figure e2. ROC Curves For Final Full and HP-HRCT Diagnosis Score-Based Models in UMHS and LTRC Cohorts

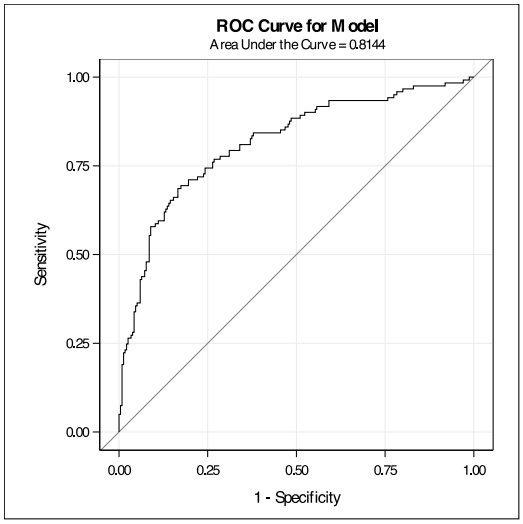
UMHS Cohort Full Model (Model 7)



LTRC Cohort Full Model



UMHS Cohort Adjusted HP-HRCT Score



LTRC Cohort Adjusted HP-HRCT Score

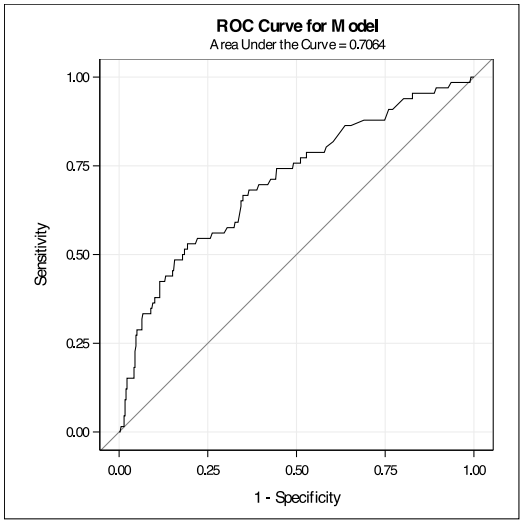


Figure e3. Calibration Curve: Predicted Probability versus Observed Frequency of HP Diagnosis in the LTRC Cohort

Figure e3. Calibration Curve: Predicted Probability versus Observed Frequency of HP Diagnosis in the LTRC Cohort

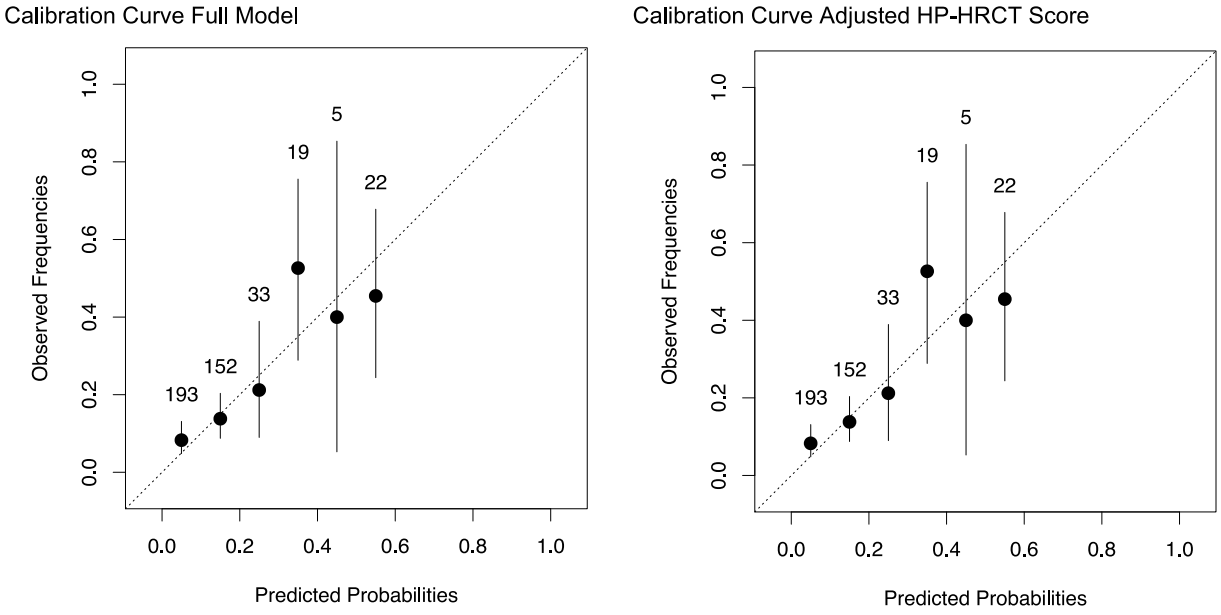


Table e8. Sensitivity Analyses

	"Classic HP" Surgical Biopsy (n=64) versus All Not-HP (n=235)				**"Classic HP" (n=64) vs Idiopathic Interstitial Pneumonias* (n=203)**				**Subjects without Exposure History (HP n=38, Not-HP n=179)**			
	Model AUC = 0.803				Model AUC = 0.818				Model AUC = 0.835			
Variable	*OR (CI 95%)*			*P*	*OR (CI 95%)*			*P*	*OR (CI 95%)*			*P*
Age	1.00 (0.97-1.03)			0.90	0.99 (0.96-1.02)			0.46	1.03 (0.99-1.07)			0.10
Male	0.48 (0.25-0.93)			0.03	0.54 (0.27-1.07)			0.08	0.42 (0.16-1.06)			0.06
Ever Smoker	0.71 (0.38-1.34)			0.29	0.68 (0.35-1.42)			0.26	0.69 (0.30-1.58)			0.38
MA-AT > Reticulation	6.04 (3.06-11.93)			<0.001	6.41 (3.13-13.14)			<0.001	4.64 (1.97-10.91)			<0.001
Axial Diffuse	2.18 (1.09-4.36)			0.03	2.19 (1.06-4.50)			0.03	5.13 (2.08-12.67)			<0.001
	HP-HRCT Diagnosis Score Test Characteristics											
Points Threshold	*HP (n)*	*Not HP (n)*	*Sens*	*Spec*	*HP (n)*	*Not HP (n)*	*Sens*	*Spec*	*HP (n)*	*Not HP (n)*	*Sens*	*Spec*
>0	50	77	78.1	67.2	50	60	78.1	70.4	33	62	86.8	65.4
>1	44	45	68.8	80.9	44	37	68.8	82.3	26	35	68.4	80.4
>2	32	23	50.0	90.2	32	19	50.0	90.6	20	17	52.6	90.5

*Idiopathic interstitial pneumonias included acute interstitial pneumonitis, cryptogenic organizing pneumonia, desquamative interstitial pneumonia, idiopathic pulmonary fibrosis, nonspecific interstitial pneumonitis, and respiratory bronchiolitis-ILD.

Table e9. "Classic HP" Characteristics (n=64)

Age, y, mean (SD)	59.5 (10.1)
Sex, No. (%) male	20 (31.3)
Smoking Status, No. (%)	
Current	3 (4.7%)
Former	27 (42.2)
Never	34 (53.1)
Honeycombing Present	4 (6.3%)
Honeycombing Semi-Quant*	0.16 (0.45)
Reticulation Present	32 (50%)
Reticulation SQ*	0.69 (0.48)
Ground Glass Present	50 (78.1%)
Ground Glass SQ*	1.68 (1.06)
Mosaic Attenuation Present	31 (48.4%)
Mosaic Attenuation SQ*	0.97 (1.04)
Air Trapping Present (n=53 with expiratory)	39 (73.6%)
Air Trapping SQ*	1.50 (0.91)
Mosaic Attenuation OR Air Trapping Present	43 (67.2%)
GG > Reticulation Present	51 (79.7%)
MA-AT > Reticulation Present	44 (68.8%)
Centrilobular Nodules Present	5 (7.8%)
Centrilobular Nodules SQ*	0.43 (0.90)
Traction Bronchiectasis Present	36 (56.3%)
Traction Bronchiectasis SQ*	2.05 (1.74)
Craniocaudal Distribution	
Upper Lung	0.19 (0.35)
Lower Lung	0.36 (0.39)
Diffuse	0.45 (0.37)
Axial Distribution	
Central	0.08 (0.17)
Peripheral	0.31 (0.36)
Subpleural	0.02 (0.08)
Peribronchovascular	0.01 (0.04)
Diffuse	0.59 (0.38)

Table e10. Positive Predictive Value of Alternative Adjusted HP Probability Based on Regression Formula

	UMHS HP Prevalence = 33.9%				LTRC HP Prevalence = 15.6%			
HP Probability	*HP (n)*	*Not-HP (n)*	*PPV (95% CI)*	*NPV (95% CI)*	*HP (n)*	*Not-HP (N)*	*PPV (95% CI)*	*NPV (95% CI)*
>/= 10%	113	169	0.40 (0.34 - 0.46)	0.89 (0.82-0.96)	53	234	0.19 (0.14 - 0.23)	0.91 (0.86-0.95)
>/= 20%	98	79	0.55 (0.48 - 0.63)	0.87 (0.82-0.92)	34	75	0.31 (0.23 - 0.40)	0.90 (0.87-0.93)
>/= 30%	91	62	0.60 (0.52 - 0.67)	0.85 (0.80-0.90)	29	49	0.37 (0.27 - 0.48)	0.89 (0.86-0.93)
>/= 40%	83	41	0.67 (0.59 - 0.75)	0.84 (0.79-0.88)	22	27	0.45 (0.31 - 0.59)	0.88 (0.85-0.92)
>/= 50%	79	35	0.69 (0.61 - 0.78)	0.83 (0.78-0.87)	22	24	0.48 (0.33 - 0.62)	0.88 (0.85-0.92)
>/= 60%	65	20	0.77 (0.68 - 0.86)	0.79 (0.75-0.84)	14	18	0.44 (0.27 - 0.61)	0.87 (0.83-0.90)
>/= 70%	47	14	0.77 (0.67 - 0.88)	0.75 (0.70-0.80)	10	12	0.46 (0.25 - 0.66)	0.86 (0.83-0.90)
>/= 80%	3	0	1.00 (1.00 - 1.00)	0.67 (0.62-0.72)	2	0	1.00 (1.00 - 1.00)	0.85 (0.81-0.88)

Adjusted HP Score = -1.0179 - 0.00639*(Age in years) – 0.6910 (if sex is male) – 0.2136 (if patient has ever smoked) + 1.8251 (if mosaic attenuation or air trapping > Reticulation) + 0.8463 (if interstitial disease has diffuse axial distribution)

Adjusted Model-based HP Probability = $e^{(\text{Adjusted HP Score})}/(1+ e^{(\text{Adjusted HP Score})})$

PPV: Positive Predictive Value; NPV: Negative Predictive Value; N: Number with HP or Not-HP