

The use of pretest probability increases the value of high-resolution CT in diagnosing usual interstitial pneumonia

Robert Brownell,¹ Teng Moua,² Travis S Henry,³ Brett M Elicker,³ Darin White,⁴ Eric Vittinghoff,⁵ Kirk D Jones,⁶ Anatoly Urisman,⁶ Carlos Aravena,⁷ Kerri A Johannson,⁸ Jeffrey A Golden,¹ Talmadge E King Jr,¹ Paul J Wolters,¹ Harold R Collard,¹ Brett Ley¹

► Additional material is published online only. To view please visit the journal online (<http://dx.doi.org/10.1136/thoraxjnl-2016-209671>).

For numbered affiliations see end of article.

Correspondence to

Dr Brett Ley, Department of Medicine, University of California San Francisco, 505 Parnassus Ave, Box 0111, San Francisco, CA 94143, USA; Brett.Ley@UCSF.edu

Received 2 November 2016
Revised 30 November 2016
Accepted 5 December 2016
Published Online First
12 January 2017

ABSTRACT

Background Recent studies have suggested that non-definitive patterns on high-resolution CT (HRCT) scan provide sufficient diagnostic specificity to forgo surgical lung biopsy in the diagnosis of idiopathic pulmonary fibrosis (IPF). The objective of this study was to determine test characteristics of non-definitive HRCT patterns for identifying histopathological usual interstitial pneumonia (UIP).

Methods Patients with biopsy-proven interstitial lung disease (ILD) and non-definitive HRCT scans were identified from two academic ILD centres. Test characteristics for HRCT patterns as predictors of UIP on surgical lung biopsy were derived and validated in independent cohorts.

Results In the derivation cohort, 64/385 (17%) had possible UIP pattern on HRCT; 321/385 (83%) had inconsistent with UIP pattern. 113/385 (29%) patients had histopathological UIP pattern in the derivation cohort. Possible UIP pattern had a specificity of 91.2% (95% CI 87.2% to 94.3%) and a positive predictive value (PPV) of 62.5% (95% CI 49.5% to 74.3%) for UIP pattern on surgical lung biopsy. The addition of age, sex and total traction bronchiectasis score improved the PPV. Inconsistent with UIP pattern demonstrated poor PPV (22.7%, 95% CI 18.3% to 27.7%). HRCT pattern specificity was nearly identical in the validation cohort (92.7%, 95% CI 82.4% to 98.0%). The substantially higher prevalence of UIP pattern in the validation cohort improved the PPV of HRCT patterns.

Conclusions A possible UIP pattern on HRCT has high specificity for UIP on surgical lung biopsy, but PPV is highly dependent on underlying prevalence. Adding clinical and radiographic features to possible UIP pattern on HRCT may provide sufficient probability of histopathological UIP across prevalence ranges to change clinical decision-making.

INTRODUCTION

Idiopathic pulmonary fibrosis (IPF) is a progressive fibrosing interstitial lung disease (ILD) of unknown aetiology defined by the presence of usual interstitial pneumonia (UIP) pattern on lung biopsy.¹ High-resolution CT (HRCT) has proven useful as a surrogate for surgical lung biopsy-proven UIP pattern, and HRCT is now accepted by the

Key messages

What is the key question?

- What are the test characteristics of non-definitive high-resolution CT patterns for histopathological usual interstitial pneumonia?

What is the bottom line?

- This study demonstrates that a possible usual interstitial pneumonia pattern on high-resolution CT has high specificity for histopathological usual interstitial pneumonia, but that positive predictive value is highly dependent on the underlying prevalence of histopathological usual interstitial pneumonia in the population.

Why read on?

- We show that in some settings, the inclusion of additional clinical and radiographic information to non-definitive imaging findings is required to identify groups of patients with a high probability of histopathological usual interstitial pneumonia.

international practice guidelines as equivalent to surgical biopsy in selected cases with a definite UIP pattern on HRCT. The HRCT 'definite UIP pattern' demonstrates basilar/subpleural distribution of reticulation with or without traction bronchiectasis, subpleural 'honeycomb' cysts and the absence of features inconsistent with UIP pattern.¹⁻⁷ Patients with definite UIP pattern on HRCT generally do not need to undergo surgical lung biopsy to establish the diagnosis of IPF, sparing them morbidity and mortality risk.⁸

Many patients with idiopathic ILD do not have a definite UIP pattern on HRCT.^{2 3 5} Such patients require surgical lung biopsy for a guideline-based diagnosis of IPF, but a substantial number will forgo surgery (due to severity of illness or patient preference) and remain without a definitive diagnosis.¹ These patients are variably referred to as 'unclassifiable ILD' or 'possible IPF'.⁹

Patients without a 'definite UIP' pattern may be classified as either possible UIP (requires the



► <http://dx.doi.org/10.1136/thoraxjnl-2016-209671>



CrossMark

To cite: Brownell R, Moua T, Henry TS, et al. *Thorax* 2017;**72**:424-429.



presence of the same criteria for definite UIP but without features of honeycombing) or inconsistent with UIP (requires the presence of features inconsistent with UIP pattern).¹ There is uncertainty surrounding the appropriate management of these patients.⁹

Recent studies have looked at the positive predictive value (PPV) of non-definitive patterns on HRCT for a diagnosis of IPF, finding them to be highly predictive in patients with idiopathic disease.^{10–12} However, these studies have been performed in selected populations of patients with high (60–100%) prevalence of IPF, which may introduce spectrum bias and inflated test characteristics.^{10–13}

Using a well defined, representative cohort of patients with biopsy-proven ILD without definite UIP pattern on HRCT, our objectives were to (1) determine the test characteristics of non-definitive HRCT patterns for histopathological UIP; (2) explore whether additional clinical and radiographic characteristics could improve these test characteristics; and (3) develop and validate a diagnostic model to aid in the clinical and radiographic identification of histopathological UIP pattern in patients with non-definitive HRCTs.

MATERIAL AND METHODS

Study population

Patient data were obtained from cohorts at the University of California, San Francisco (UCSF, derivation cohort) and the Mayo Clinic, Rochester (validation cohort). The source population for the derivation cohort included all patients evaluated at the UCSF ILD clinic from September 2002 to July 2015, with all clinical diagnoses established through multidisciplinary team discussion.¹ Inclusion criteria were availability of prospectively scored surgical lung biopsy and an HRCT scan within 1 year of the biopsy. Patients with a definite UIP pattern on HRCT, a diagnosis of connective tissue disease at the time of biopsy or a predominately cystic lung disease were excluded from the primary analysis cohort. The validation cohort included patients evaluated at the Mayo Clinic, Rochester between December 1999 and February 2016 with a pathological diagnosis of fibrotic ILD (IPF, non-specific interstitial pneumonia, hypersensitivity pneumonia, desquamative interstitial pneumonia and unclassifiable or undifferentiated interstitial fibrosis). Inclusion and exclusion criteria were otherwise identical.

Radiological and clinical predictor variables

All eligible HRCT scans were deidentified. HRCTs in the UCSF cohort were reviewed independently by two experienced chest radiologists (BME and TSH), and HRCTs in the Mayo Clinic cohort were reviewed by a third expert chest radiologist (DW). Radiologists were blinded to histopathology results and clinical information. Published criteria were used to categorise HRCT scans as definite UIP, possible UIP or inconsistent with UIP pattern.¹ For HRCT scans categorised as possible UIP pattern, the extent of radiographic traction bronchiectasis was scored separately in each lobe (right upper, middle, and lower lobes and left upper, lingula, and lower lobes) as 0-absent, 1-mild, 2-moderate or 3-severe, and then summed to provide a total traction bronchiectasis score as previously described (for CT image examples see online supplementary figure S1).^{14–16} For HRCT scans categorised as inconsistent with UIP pattern, the inconsistent features were recorded as binary variables (yes/no). These features were upper/mid-lung predominance, peribronchovascular predominance, extensive ground glass opacities, profuse micronodules, discrete cysts in areas away from honeycombing, mosaic attenuation/air trapping in three or more lobes and

consolidations.^{1 17} Clinical variables included were age, sex and smoking history (ever vs never smoker).

Histopathological diagnosis

All surgical lung biopsies were reviewed and interpreted prospectively by expert pulmonary pathologists and given a final histopathological diagnosis. For the purposes of this study, final histopathological diagnoses were retrospectively categorised as definite/probable UIP, possible UIP/unclassifiable pulmonary fibrosis (PF), or not UIP in concordance with published diagnostic criteria (see online supplementary table S1).¹ The outcome ‘histopathologic UIP’ was considered positive for ‘definite/probable UIP’ and negative for ‘possible UIP/unclassifiable PF’ or ‘not UIP’.

Statistical analysis

Test characteristics with 95% CIs were determined for possible UIP pattern on HRCT (alone and with additional covariates of age, sex and traction bronchiectasis score) and inconsistent with UIP pattern in both cohorts. Test characteristics included sensitivity, specificity, PPV, negative predictive value, area under the receiver-operator curve and positive likelihood ratios (LR+). To further demonstrate the effect of differing population prevalence rates (ie, differing pretest probabilities) of histopathological UIP on post-test probabilities, histopathological UIP prevalence rates were varied across a hypothetical range.^{10 12} Several sensitivity analyses were performed: including cases with possible UIP/unclassifiable PF on histopathology as ‘positive’ outcome (ie, as equivalent to definite/probable UIP cases); using the clinical diagnosis of IPF rather than histopathological UIP as the outcome; including patients with diagnoses of a defined connective tissue disease (CTD) prior to biopsy; and using individual radiologist interpretations of HRCT pattern rather than consensus pattern.

To evaluate the added value of clinical and radiographic features for predicting histopathological UIP using non-diagnostic HRCT, the least absolute shrinkage and selection operator (LASSO) method¹⁸ was used to select the most important predictors among candidates. Candidate predictor variables included were age (categorically defined as ages ≥ 50 years, 60 years and 70 years), sex, smoking history (ever smoker vs never smoker), HRCT possible UIP pattern and HRCT possible UIP pattern with a total traction bronchiectasis score ≥ 4 , (which was near the median score). Once key predictor variables were identified by LASSO, model coefficients were rescaled to develop a simplified point-score model (the UIP score model). Model discrimination was evaluated by the C-statistic, and interval test characteristics were determined for each point score level. The UIP score model was then validated in the Mayo Clinic cohort. In an effort to identify radiographic features predictive of non-UIP histopathology, a second model was constructed in the subgroup with HRCTs that were inconsistent with UIP. Candidate predictor variables were each individual inconsistent HRCT feature, the presence of HRCT fibrotic features (reticulation, traction bronchiectasis and honeycombing), age, sex and smoking history.

RESULTS

Patient characteristics

Four hundred and fifty-five patients with prospectively scored surgical lung biopsies and HRCT scans were identified in the derivation cohort and included in the study. Of these, 49 had a definite UIP pattern on HRCT and 21 had documented CTD at the time of surgical lung biopsy. This left a primary study cohort of 385 patients (see online supplementary figure S2).

Table 1 Cohort characteristics by HRCT pattern

Characteristic	HRCT		HRCT	
	Possible UIP		Inconsistent with UIP	
	Derivation (n=64)	Validation (n=71)	Derivation (n=321)	Validation (n=95)
Age in years, mean (SD)	65 (8)	68 (8)	59 (12)	61 (12)
Male sex, n (%)	40 (62.5)	47 (66.0)	134 (41.7)	50 (53.0)
Ever smoker, n (%)	36 (56.2)	42 (59.0)	174 (54.2)	47 (49.0)
Race/ethnicity, n (%)				
White or Caucasian	48 (75.0)	69 (97.0)	238 (74.1)	90 (95.0)
Hispanic or Latino	9 (14.1)	2 (3.0)	31 (9.7)	2 (2.0)
Asian	5 (7.8)	0	16 (5.0)	1 (1.0)
Black or African American	2 (3.1)	0	14 (4.4)	1 (1.0)
Native American	0	0	8 (2.5)	0
Other	0	0	7 (2.2)	1 (1.0)
Unknown/declined	0	0	7 (2.2)	0
PFTs, mean (SD)				
FVC, % predicted	70.8 (18.9)	69.1 (16.2)	70.6 (18.9)	61.9 (13.8)
FEV ₁ , % predicted	75.7 (19.4)	74.2 (15.7)	75.4 (19.7)	64.8 (14.7)
TLC, % predicted	74.0 (17.1)	68.1 (12.1)	74.3 (17.6)	67.2 (12.3)
DLCO, % predicted	53.1 (19.0)	47.5 (14.3)	53.5 (19.4)	46.3 (12.1)
Clinical diagnosis, n (%)				
IPF	39 (60.9)	69 (97.0)	57 (17.8)	49 (46.0)
HP	12 (18.8)	2 (3.0)	89 (27.7)	25 (26.0)
CTD-ILD*	3 (4.7)	0	17 (5.3)	4 (4.0)
Idiopathic NSIP	1 (1.6)	0	25 (7.8)	7 (7.0)
Undifferentiated CTD	2 (3.1)	0	23 (7.2)	0
Unclassifiable ILD	7 (10.9)	0	52 (16.2)	6 (6.3)
Other†	0	0	58 (18.1)	4 (4.2)

*CTD-ILD diagnosed after surgical lung biopsy included:

UCSF cohort: Anti-neutrophil cytoplasmic antibody-associated vasculitis (n=1), inflammatory myositis (n=9), mixed connective tissue disease (n=1), rheumatoid arthritis (n=3), systemic lupus erythematosus (n=1) and scleroderma (n=4). *Mayo cohort:* Sjogren's syndrome (n=1), inflammatory myositis (n=1), mixed connective tissue disease (n=1), scleroderma (n=1).

†Other diagnoses included:

UCSF cohort: Asbestosis (n=1), bronchiolitis (n=8), cryptogenic organising pneumonia (n=7), diffuse alveolar haemorrhage (n=1), desquamative interstitial pneumonia (n=8), dyskeratosis congenita (n=1), eosinophilic pneumonia (n=2), familial pulmonary fibrosis (n=1), mycobacterium avium complex (n=2), lymphocytic interstitial pneumonia (n=3), methotrexate toxicity (n=1), nitrofurantoin toxicity (n=1), not ILD (n=1), surfactant protein C mutation (n=1), light chain deposition disease (n=1), lipoid pneumonia (n=1), respiratory bronchiolitis ILD (n=3), sarcoidosis (n=15).

Mayo Cohort: Desquamative interstitial pneumonia (n=4).

CTD, connective tissue disease; HP, hypersensitivity pneumonitis; HRCT, high-resolution CT; ILD, interstitial lung disease; IPF, idiopathic pulmonary fibrosis; DLCO, diffusing capacity for carbon monoxide; NSIP, non-specific interstitial pneumonitis; PFTs, pulmonary function tests; TLC, total lung capacity; UCSF, University of California, San Francisco; UIP, usual interstitial pneumonia.

Cohort characteristics are summarised in [table 1](#). Sixty-four patients (17%) had possible UIP pattern on HRCT; 321 (83%) had inconsistent with UIP pattern on HRCT. There was good agreement between radiologists for categorising HRCT patterns as definite, possible or inconsistent with UIP ($\kappa=0.73$; 95% CI 0.65 to 0.82). The most common clinical diagnosis in the derivation cohort was chronic hypersensitivity pneumonitis (HP) (n=101/385, 26%) and IPF (n=96/385, 25%). The prevalence of definite/probable UIP on surgical lung biopsy was 29% (113/385) (see online supplementary table S1 for other pathology designations).

Test characteristics of non-definitive CT patterns

A possible UIP pattern on HRCT demonstrated a specificity of 91.2% (95% CI 87.2% to 94.3%), a PPV of 62.5% (95% CI 49.5% to 74.3%) and an LR+ of 4.01 (95% CI 2.54 to 6.33) for definite/probable UIP on surgical biopsy ([table 2](#)). The most common alternative histopathological interpretations among patients that had a possible UIP pattern on HRCT but did not have definite/probable UIP on histopathology (n=24) included fibrotic non-specific interstitial pneumonia (NSIP) with or

without interstitial granulomas (n=9/24), prominent bronchio-centric fibrosis with or without interstitial granulomas (n=5/24) and unclassifiable fibrosing interstitial pneumonia (n=6/24) (see online supplementary table S2). A possible UIP pattern on HRCT along with a total traction bronchiectasis score ≥ 4 , increased age (ie, age ≥ 60 years) and male sex all increased the PPV and LR+ for definite/probable UIP on surgical biopsy, but at the expense of reduced sensitivity ([table 2](#)).

Among patients who had HRCT scans inconsistent with UIP, 22.7% had definite/probable UIP on biopsy ([table 2](#)). The most common radiological features in patients with HRCT scans that were inconsistent with UIP were peribronchovascular distribution (44%), ground glass opacities (39%), mosaic perfusion/air trapping (23%) and upper-mid lung predominance (16%). The inconsistent features least associated with definite/probable UIP on biopsy were diffuse micronodules, consolidation and ground glass opacities (see online supplementary table S3).

Prediction of histopathological UIP pattern

In evaluating multivariate prediction models for histopathological UIP, the LASSO procedure selected the variables age

Table 2 Test characteristics of HRCT possible UIP pattern, alone and with consideration of other clinical and radiographic features, for definite/probable histopathological UIP

Pattern	Sensitivity	Specificity	PPV	NPV	AUROC	LR+
Possible UIP-UCSF	35.4 (26.6–45.0)	91.2 (87.2–94.3)	62.5 (49.5–74.3)	77.3 (72.3–81.7)	0.63 (0.58–0.68)	4.01 (2.54–6.33)
Possible UIP-Mayo	60.4 (50.6–69.5)	92.7 (82.4–98.0)	94.4 (86.2–98.4)	53.7 (43.2–64.0)	0.77 (0.71–0.82)	8.30 (3.19–21.58)
Possible UIP with traction bronchiectasis score $\geq 4^*$ and male ≥ 60 -UCSF	16.8 (10.4–25.0)	99.6 (98.0–100)	95.0 (75.1–99.9)	74.2 (69.4–78.7)	0.58 (0.55–0.62)	45.7 (6.2–338)
Possible UIP with traction bronchiectasis score $\geq 4^*$ and male ≥ 60 -Mayo	32.4 (23.9–42.0)	98.2 (90.3–100.0)	97.3 (85.8–99.9)	41.9 (33.2–50.9)	0.65 (0.61–0.70)	17.84 (2.51–126.70)
Inconsistent with UIP-UCSF	64.6 (55.0–73.4)	8.8 (5.7–12.8)	22.7 (18.3–27.7)	37.5 (25.7–50.5)	0.37 (0.32–0.41)	0.71 (0.61–0.82)
Inconsistent with UIP-Mayo	39.6 (30.5–49.4)	7.3 (2.0–17.6)	46.3 (36.0–56.8)	5.6 (1.6–13.8)	0.23 (0.18–0.29)	0.43 (0.34–0.54)

*The extent of radiographic traction bronchiectasis was scored in each lobe (right upper, middle, and lower lobes and left upper, lingula, and lower lobes) as 0-absent, 1-mild, 2-moderate, or 3-severe, and then summed to provide a total traction bronchiectasis score (see online supplementary figure S1 and ref. 14). AUROC, area under the receiver-operator curve; HRCT, high-resolution CT; LR+, positive likelihood ratio; NPV, negative predictive value; PPV, positive predictive value; UCSF, University of California, San Francisco; UIP, usual interstitial pneumonia.

Table 3 The usual interstitial pneumonia (UIP) score model

Characteristic	Points
Age, in years	
50–59	2
≥ 60	3
Male sex	1
Possible UIP+total traction bronchiectasis score ≥ 4	6
Total score possible	10

The extent of radiographic traction bronchiectasis was scored in each lobe (right upper, middle, and lower lobes and left upper, lingula, and lower lobes) as 0-absent, 1-mild, 2-moderate, or 3-severe, and then summed to provide a total traction bronchiectasis score (see online supplementary figure S1 and ref. 14).

(50–59 years or ≥ 60 years), male sex and HRCT possible UIP pattern with a total traction bronchiectasis score ≥ 4 . A simple point-score extension of this model (the UIP score) is shown in table 3 (also see online supplementary table S4). The UIP score had good discriminative performance for histopathological UIP (C statistic 0.74, 95% CI 0.69 to 0.78). Models constructed in the subgroup with inconsistent with UIP pattern on HRCT (ie, excluding patients with possible UIP pattern on HRCT) did not provide sufficient predictive performance to rule in histopathological UIP for any combination of clinical and radiographic features (maximum model-estimated PPV 38%).

Sensitivity analyses

In sensitivity analyses, counting cases with histopathological possible UIP/unclassifiable PF as equivalent to definite/probable UIP in the outcome of ‘histopathologic UIP’ substantially increased the PPV/LR+ for possible UIP pattern on HRCT (see online supplementary table S5). We observed no improvement in PPV for possible UIP pattern on HRCT for the predicted outcome of a clinical diagnosis of IPF (rather than histopathological UIP) or by including those with a diagnosis of defined CTD prior to biopsy (see online supplementary table S5). Finally, results were similar when conducted by individual radiologist interpretation of HRCT pattern rather than by consensus pattern (see online supplementary table S6) in our derivation cohort.

Validation and generalisability

The validation cohort contained 166 patients; the prevalence of definite/probable UIP on surgical lung biopsy was 67% (111/166) (see online supplementary figure S3). A possible UIP pattern on HRCT demonstrated similar specificity to the

derivation cohort (92.7%, 95% CI 82.4 to 98.0) (table 2). Because of the much higher population prevalence of histopathological UIP (67% vs 29%), the PPV was greatly increased (94.4%, 95% CI 86.2 to 98.4). The UIP score maintained good discriminative performance for histopathological UIP (C statistic 0.83, 95% CI 0.76 to 0.89).

The relationship of population prevalence of histopathological UIP to PPV of the UIP score is shown in figure 1, based on test characteristics (sensitivity and specificity) from the derivation cohort. For any given UIP score, PPV increases with increasing prevalence. For example, a UIP score of 4 has a PPV of 60% in a population with 40% prevalence of histopathological UIP but a PPV $> 80\%$ in a population with 70% prevalence of histopathological UIP.

DISCUSSION

In this study, we demonstrate that a possible UIP pattern on HRCT has high specificity for histopathological UIP, but that PPV is highly dependent on the clinician’s pretest probability that the biopsy will show UIP. Fundamental components of pretest probability include individual patient information and disease prevalence in the relevant population. In high prevalence settings, a possible UIP pattern on HRCT may be sufficiently predictive of histopathological UIP to forgo a surgical lung biopsy. In lower prevalence settings, however, the inclusion of additional information (ie, age, sex and traction bronchiectasis score) is required to identify groups of patients with a high PPV for histopathological UIP.

As a predictor of histopathological UIP, HRCT pattern is used as a diagnostic test. Applying a diagnostic test to clinical decision-making requires the integration of three key pieces of information: (1) the threshold probability of the outcome at which the clinician would change his or her next-step action (eg, proceed to surgical lung biopsy or not); (2) the pretest probability of the outcome (often equal to the population prevalence); and (3) the diagnostic test’s characteristics. We discuss each of these three components below.

Threshold probability

The threshold probability of UIP on biopsy, at which clinicians would be comfortable assuming the presence of histopathological UIP without proceeding to surgical lung biopsy, has not been clearly established for patients with suspected IPF. However, a definite UIP pattern on HRCT, with an estimated PPV of 90% for histopathological UIP, has been widely accepted as sufficiently predictive to serve as a surrogate for surgical lung

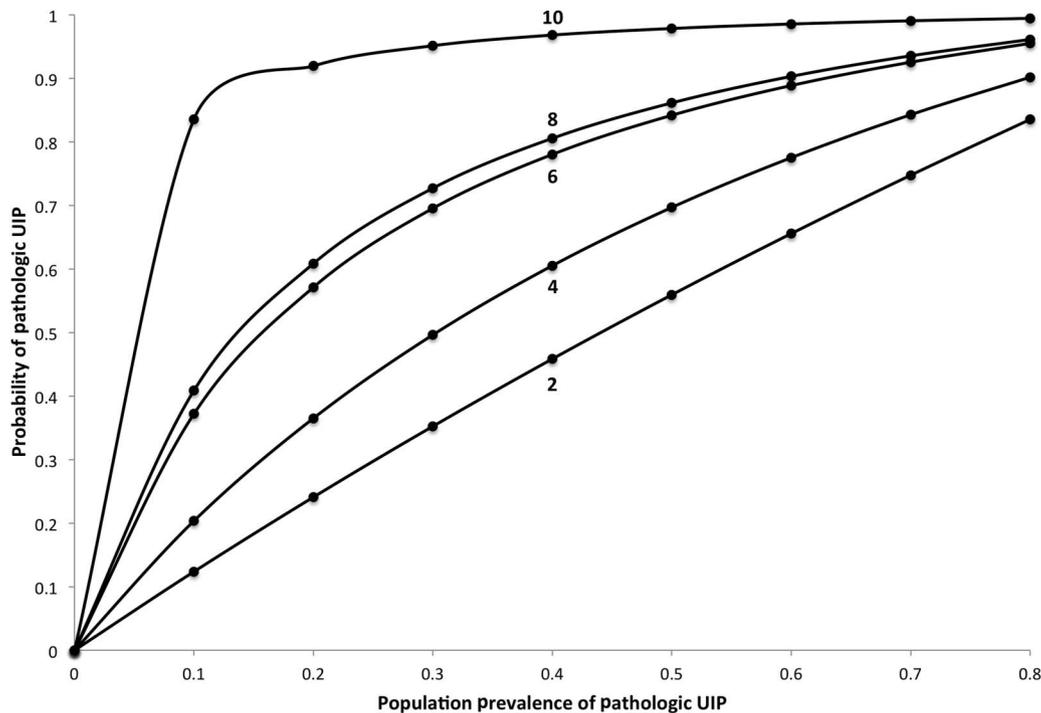


Figure 1 Plot demonstrates the probability of histopathological usual interstitial pneumonia (UIP), based on the interval likelihood ratios (LRs) from the University of California, San Francisco cohort, across hypothetical population prevalence rates of histopathological UIP for individual point levels of the UIP score model calculated from a patient's age (+2 points for age 50–59 years or +3 points for age ≥60 years), male sex (+1 points) and high-resolution CT possible UIP pattern with a total traction bronchiectasis score ≥4 (+6 points) for a total of 10 possible points.

biopsy.^{2–5} It therefore seems reasonable to assume this threshold probability for the purposes of further discussion.

Pretest probability of histopathological UIP

Our derivation and validation cohorts had widely differing population prevalences (pretest probabilities) of histopathological UIP. Previously published cohorts have generally been high.^{10–12, 19} There are likely several reasons for these differences including geographical variation in disease frequency (the high prevalence of chronic HP in the UCSF cohort), inter-rater differences among pathologists, and selection or spectrum bias. In contrast to previous studies, we included all patients undergoing surgical lung biopsy for workup of ILD, and in the case of the UCSF cohort, without any preselection by clinical diagnosis. Previous studies have been enriched for patients with IPF (and therefore histopathological UIP) due to including patients enrolled in IPF clinical trials^{10, 13} and/or excluding subjects based on other common clinical diagnoses.^{12, 13}

Test characteristics of diagnostic test

Our data show that the PPV and LR+ of possible UIP is insufficient in some populations to push the probability of histopathological UIP above the threshold of 90%. Unless the population prevalence is 70% or higher, a more accurate diagnostic test is required. We identified two key clinical predictors (patient age and sex) and one additional radiographic predictor (total traction bronchiectasis score) that significantly increased the PPV of possible UIP on HRCT for histopathological UIP, and combined these predictors into the UIP score. Among patients with an inconsistent with UIP pattern on HRCT, we were unable to identify any combination of clinical and radiographic features with sufficiently high PPV to exceed the diagnostic threshold for histopathological UIP.

It has been previously observed that older age predicts a clinical diagnosis of biopsy-proven IPF among patients with fibrosis on HRCT in a population with high IPF prevalence (70%).¹² While the presence of traction bronchiectasis in patients with possible UIP pattern on HRCT has not been previously identified as a predictor of histopathological UIP pattern, it has been used as criteria for inclusion in IPF trials,²⁰ and has previously been associated with poor outcomes in patients with fibrotic ILD.^{15, 16} The UIP score model can thus be used in areas of low or unknown UIP prevalence to identify patients with sufficiently high post-test probability of histopathological UIP that clinicians would feel comfortable avoiding a surgical biopsy.

The lesson from clinical decision-making is that simply applying the PPV of a possible UIP pattern on HRCT without regard for the population prevalence in which it was derived and to which it is being applied risks drastically overestimating (or underestimating) the true PPV of the test in the patient under consideration. We recognise that many providers will not have detailed information on the local prevalence of histopathological UIP among patients with ILD with non-definitive HRCTs. In these situations, we suggest that the local prevalence of IPF among patients with ILD with non-definitive HRCTs may closely approximate that of histopathological UIP, as it did in both of our study populations and could be substituted. However, this should be considered with caution in centres that have not routinely used surgical lung biopsy in the diagnostic workup of patients with non-definitive HRCT scans, since they may be unaware of the rates of competing histopathological diagnoses.

Our study has several strengths, most critically the inclusion of a diverse group of unselected and well characterised patients undergoing surgical lung biopsy as part of their clinical evaluation for ILD. All patients had detailed clinical, radiological and

pathological data available for this study, and all patients received a multidisciplinary diagnosis supported by clinical guidelines. A second strength is that we considered histopathological diagnosis as our outcome, rather than clinical diagnosis. We believe this is the more relevant outcome, and it avoids the issue of confirmation bias, in which the diagnostic test being studied informs the outcome of interest. Finally, we approached the diagnostic test analyses comprehensively, incorporating the effects of pretest probability, test characteristics and threshold limits to fully describe the use of HRCT pattern in clinical practice.

Due to the limitations of our retrospective study design, we were unable to fully account for other clinical variables that may affect pretest probability of histopathological UIP such as symptoms and serologies suggestive of autoimmune disease and exposures associated with HP. However, we believe that cases selected by treating ILD clinicians to undergo surgical lung biopsy presumably lacked adequate clinical evidence for an alternative diagnosis. Other limitations relate to generalisability (both study cohorts are tertiary referral centres and may be more likely to have atypical cases), and the unavailability of biopsy slides for re-review by multiple pathologists (allowing us to evaluate the sensitivity of test characteristics to interobserver variation in surgical lung biopsy interpretation). Systematic differences in pathologist interpretations (eg, one pathologist being systematically more or less likely to call UIP compared with another pathologist) would be expected to bias our results.

In conclusion, we demonstrate that the use of non-definitive patterns on HRCT as surrogates for surgical lung biopsy requires careful consideration of the principles of diagnostic tests and clinical decision-making, of which population prevalence (ie, pretest probability) is a central component. HRCT pattern alone (either possible UIP pattern or inconsistent with UIP pattern) is insufficiently predictive in lower UIP prevalence populations. The addition of clinical and radiological features through the use of the UIP score model may allow clinicians and clinical trialists to identify patients at sufficiently high probability of histopathological UIP pattern that surgical biopsy can be avoided. Indeed, it may be that IPF can be diagnosed in these cases if no alternative aetiology for histopathological UIP is identified. Similar logic supports the clinical diagnosis of IPF in patients with a definite UIP pattern on HRCT. Such a change in practice will need to await careful review by the ILD community.

Author affiliations

¹Department of Medicine, University of California San Francisco, San Francisco, California, USA

²Department of Medicine, Mayo Clinic, Rochester, Minnesota, USA

³Department of Radiology and Biomedical Imaging, University of California San Francisco, San Francisco, California, USA

⁴Department of Radiology, Mayo Clinic, Rochester, Minnesota, USA

⁵Department of Epidemiology and Biostatistics, University of California San Francisco, San Francisco, California, USA

⁶Department of Pathology, University of California San Francisco, San Francisco, California, USA

⁷Respiratory Diseases Department, Pontifical Catholic University, Santiago, Chile

⁸Department of Medicine, University of Calgary, Calgary, Canada

Contributors RB, HRC and BL designed the study. RB, BL, TSH, BME, KDJ, AU, CA, KAJ, TM and DW collected data. RB, BL and EV performed the data analysis.

All authors contributed to data interpretation. RB, BL and HRC wrote the original manuscript. All authors contributed to revisions of the manuscript, provided final approval of the version to be published and agree to be accountable for all aspects of the work.

Funding Research reported in this publication was supported by the National Center for Advancing Translational Sciences of the NIH under Award Number KL2TR001870 as well as NIH/NHLBI grants F32HL124895 and K24HL127131. The Nina Ireland Program for Lung Health supports the UCSF ILD clinical database.

Competing interests None declared.

Ethics approval IRB at UCSF and Mayo Clinic, Rochester.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

- 1 Raghu G, Collard HR, Egan JJ, *et al.* An official ATS/ERS/JRS/ALAT statement: idiopathic pulmonary fibrosis: evidence-based guidelines for diagnosis and management. *Am J Respir Crit Care Med* 2011;183:788–824.
- 2 Hunninghake GW, Lynch DA, Galvin JR, *et al.* Radiologic findings are strongly associated with a pathologic diagnosis of usual interstitial pneumonia. *Chest* 2003;124:1215–23.
- 3 Hunninghake GW, Zimmerman MB, Schwartz DA, *et al.* Utility of a lung biopsy for the diagnosis of idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2001;164:193–6.
- 4 Quadrelli S, Molinari L, Ciallella L, *et al.* Radiological versus histopathological diagnosis of usual interstitial pneumonia in the clinical practice: does it have any survival difference? *Respiration* 2010;79:32–7.
- 5 Raghu G, Mageto YN, Lockhart D, *et al.* The accuracy of the clinical diagnosis of new-onset idiopathic pulmonary fibrosis and other interstitial lung disease: a prospective study. *Chest* 1999;116:1168–74.
- 6 Swensen SJ, Aughenbaugh GL, Myers JL. Diffuse lung disease: diagnostic accuracy of CT in patients undergoing surgical biopsy of the lung. *Radiology* 1997;205:229–34.
- 7 Lynch DA, Godwin JD, Safrin S, *et al.* High-resolution computed tomography in idiopathic pulmonary fibrosis: diagnosis and prognosis. *Am J Respir Crit Care Med* 2005;172:488–93.
- 8 Hutchinson JP, Fogarty AW, McKeever TM, *et al.* In-hospital mortality following surgical lung biopsy for interstitial lung disease in the USA: 2000–2011. *Am J Respir Crit Care Med* 2016;193:1161–7.
- 9 Ryerson CJ, Urbania TH, Richeldi L, *et al.* Prevalence and prognosis of unclassifiable interstitial lung disease. *Eur Respir J* 2013;42:750–7.
- 10 Chung JH, Chawla A, Peljto AL, *et al.* CT scan findings of probable usual interstitial pneumonitis have a high predictive value for histologic usual interstitial pneumonitis. *Chest* 2015;147:450–9.
- 11 Raghu G, Lynch D, Godwin JD, *et al.* Diagnosis of idiopathic pulmonary fibrosis with high-resolution CT in patients with little or no radiological evidence of honeycombing: secondary analysis of a randomised, controlled trial. *Lancet Respir Med* 2014;2:277–84.
- 12 Fell CD, Martinez FJ, Liu LX, *et al.* Clinical predictors of a diagnosis of idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2010;181:832–7.
- 13 Yagihashi K, Huckleberry J, Colby TV, *et al.* Radiologic-pathologic discordance in biopsy-proven usual interstitial pneumonia. *Eur Respir J* 2016;47:1189–97.
- 14 Edey AJ, Devaraj AA, Barker RP, *et al.* Fibrotic idiopathic interstitial pneumonias: HRCT findings that predict mortality. *Eur Radiol* 2011;21:1586–93.
- 15 Sumikawa H, Johkoh T, Colby TV, *et al.* Computed tomography findings in pathological usual interstitial pneumonia: relationship to survival. *Am J Respir Crit Care Med* 2008;177:433–9.
- 16 Sumikawa H, Johkoh T, Ichikado K, *et al.* Usual interstitial pneumonia and chronic idiopathic interstitial pneumonia: analysis of CT appearance in 92 patients. *Radiology* 2006;241:258–66.
- 17 Hansell DM, Bankier AA, MacMahon H, *et al.* Fleischner Society: glossary of terms for thoracic imaging. *Radiology* 2008;246:697–722.
- 18 Tibshirani R. Regression shrinkage and selection via the LASSO. *J R Stat Soc Series B* 1996;58:267–88.
- 19 Salisbury ML, Xia M, Murray S, *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.
- 20 Richeldi L, du Bois RM, Raghu G, *et al.* Efficacy and safety of nintedanib in idiopathic pulmonary fibrosis. *N Engl J Med* 2014;370:2071–82.