≋CHEST®

Derivation and Validation of a Diagnostic Prediction Tool for Interstitial Lung Disease

Check for updates

Janelle Vu Pugashetti, MD; Aleksander Kitich, MD; Shehabaldin Alqalyoobi, MD; Anne-Catherine Maynard-Paquette, MD; David Pritchard, MD; Julia Graham, MD; Noelle Boctor, MD; Andrea Kulinich, BS; Elyse Lafond, MD; Elena Foster, MS; Cesar Mendez, BS; Saad Choudhry, MD; Jean Chalaoui, MD; Julie Morisset, MD; Michael Kadoch, MD; and Justin M. Oldham, MD

BACKGROUND: Interstitial lung disease (ILD) results in high morbidity and health-care utilization. Diagnostic delays remain common and often occur in nonpulmonology settings. Screening for ILD in these settings has the potential to reduce diagnostic delays and improve patient outcomes.

RESEARCH QUESTION: This study sought to determine whether a pulmonary function test (PFT)-derived diagnostic prediction tool (ILD-Screen) could accurately identify incident ILD cases in patients undergoing PFT in nonpulmonology settings.

STUDY DESIGN AND METHODS: Clinical and physiologic PFT variables predictive of ILD were identified by using iterative multivariable logistic regression models. ILD status was determined by using a multi-reader approach. An ILD-Screen score was generated by using final regression model coefficients, with a score $\geq$ 8 considered positive. ILD-Screen test performance was validated in an independent external cohort and applied prospectively to PFTs over 1 year to identify incident ILD cases at our institution.

RESULTS: Variables comprising the ILD-Screen were age, height, total lung capacity, FEV_1, diffusion capacity, and PFT indication. The ILD-Screen showed consistent test performance across cohorts, with a sensitivity of 0.79 and a specificity of 0.83 when applied prospectively. A positive ILD-Screen strongly predicted ILD (OR, 18.6; 95% CI, 9.4-36.9) and outperformed common ILD clinical features, including cough, dyspnea, lung crackles, and restrictive lung physiology. Prospective ILD-Screen application resulted in a higher proportion of patients undergoing chest CT imaging compared with a historical control cohort (74% vs 56%, respectively; $P = .003$), with a significantly shorter median time to chest CT imaging (5.6 vs 21.1 months; $P < .001$).

INTERPRETATION: The ILD-Screen showed good test performance in predicting ILD across diverse geographic settings and when applied prospectively. Systematic ILD-Screen application has the potential to reduce diagnostic delays and facilitate earlier intervention in patients with ILD.

CHEST 2020; 158(2):620-629

KEY WORDS: idiopathic pulmonary fibrosis; interstitial lung abnormalities; interstitial lung disease; pulmonary fibrosis; screening

---

FOR EDITORIAL COMMENT, SEE PAGE 444

Interstitial lung disease (ILD) comprises a heterogeneous group of diffuse parenchymal lung disorders that commonly result in pulmonary fibrosis. Those who develop pulmonary fibrosis have high health-care utilization[1,2] and mortality risk that is proportional to the extent of fibrosis.[3,4] Although studies have shown that antifibrotic and immunosuppressive therapy slow the progression of several common ILD subtypes,[5-8] prolonged delays in diagnosis remain common.[9,10] The factors influencing diagnostic delays remain incompletely characterized, but a large proportion of delays occur in primary care and other nonpulmonology settings.[11,12]

ILD screening in the general population continues to be challenging given the low disease prevalence and nonspecific symptoms. A more cost-effective approach may lie in the identification of high-risk populations in primary care and other clinical settings. One such group includes those undergoing pulmonary function tests (PFTs). This diagnostic modality, which includes measures of lung volumes, spirometry, and gas transfer, is used to evaluate pulmonary and nonpulmonary conditions across a broad range of medical specialties. PFT data help to characterize ILD severity[13-15]; characteristic abnormalities, including restrictive physiology and reduced gas transfer,[16] can also suggest the presence of ILD in undiagnosed patients.

The current study derived a diagnostic prediction tool for ILD (ILD-Screen) using clinical and physiologic measures collected during PFT. We validated the ILD-Screen in an independent cohort and prospectively applied this tool over 1 year to PFTs performed in nonpulmonology settings to identify incident ILD cases at our institution. ILD-Screen test performance was compared vs clinical and physiologic features of ILD, and the effectiveness of prospective ILD-Screen application on subsequent chest CT acquisition was assessed.

Patients and Methods

This investigation was performed at the University of California at Davis and the Centre Hospitalier de l'Université de Montréal (CHUM). It was approved by the institutional review board at each center, which provided waivers of consent (UC-Davis protocol #928979 and CHUM protocol #2019-7786).

ILD-Screen Derivation

Consecutive patients completing a PFT (plethysmography, spirometry, and gas transfer) from January 1, 2015, to December 31, 2016, were identified. Those with a chest CT scan performed within 6 months of PFT were included in the analysis. Variables captured on the PFT report, including age, sex, reference race, height, weight, PFT indication, total lung capacity (TLC), FVC, FEV_1, and diffusing capacity of the lung for carbon monoxide (D_{LCO}) (unadjusted for hemoglobin), were extracted and included in the ILD-Screen modeling. Patients undergoing PFT following lung transplantation were excluded.

A pulmonologist with ILD expertise (J. M. O.) and a general radiologist (A. K.) blinded to PFT data independently reviewed all chest CT scans to assess for the presence of ILD, defined as bilateral, nondependent reticular opacities affecting $\geq$ 5% of the total area of each lung. This definition was chosen to identify those with fibrotic ILD and nonfibrotic interstitial lung abnormalities (ILAs).[17-21] Patients with unilateral fibrosis, as is often observed in radiation-induced fibrosis, were not considered to have ILD for the purposes of this investigation. A thoracic radiologist (M. K.) with 9 years' experience adjudicated discordant cases and confirmed all positive cases.

Variable selection using the aforementioned variables was performed via iterative logistic regression models, with ILD modeled as a dichotomous dependent variable. PFT indication was modeled as a categorical variable with indications having an ILD prevalence of < 10%, 10% to 20%, 20% to 40%, and > 40% assigned a score of 0, 1, 2, and 3, respectively. Exploratory analyses were performed to determine whether variable derivatives (eg, FEV_1/FVC ratio, BMI), linear splines, and/or interaction variables improved model performance. Area under the curve (AUC) analysis was used to select the final regression model, with parsimony prioritized.

Beta-coefficients from variables comprising the final regression model were used to construct an ILD-Screen score, which ranged from 0 to 15. Test performance characteristics at various ILD-Screen score thresholds were assessed to determine the optimal score for ILD discrimination, with balanced sensitivity and specificity prioritized.

External Validation

Consecutive patients undergoing complete PFT at the CHUM from December 1, 2017, to May 31, 2018, were identified. Individuals with a paired chest CT scan performed within 6 months of PFT were included in the analysis. A chest radiologist (J. C.) with 41 years' experience determined ILD status as described earlier.

Foster; and Mr Mendez), Department of Internal Medicine (Drs Pritchard, Graham, and Boctor), and Department of Radiology (Dr Kadoch), University of California at Davis, Sacramento, CA; Department of Radiology (Dr Kitich), University of California at Los Angeles, Los Angeles, CA; Département de Médecine (Drs Maynard-Paquette and Morisset) and Département de Radiologie (Dr Chalaoui), Centre Hospitalier de l'Université de Montréal, Montreal, QC, Canada; and the Division of Pulmonary and Critical Care Medicine (Dr Lafond), Weill Cornell Medicine, New York, NY.

CORRESPONDENCE TO: Justin M. Oldham, MD, 4150 V St, Ste 3400, Sacramento, CA 95817; e-mail: joldham@ucdavis.edu

The ILD-Screen was applied on a weekly basis to PFTs performed at the University of California at Davis from January 1, 2017, to December 31, 2017. PFTs were excluded when: (1) ordered by a pulmonologist; (2) included in the derivation cohort; (3) ordered for a pregnant female subject; or (4) ordered for an individual aged < 18 years. To identify only incident cases, ILD-Screen-positive patients with previous chest CT imaging were excluded. The PFT-ordering physician for patients with a positive ILD-Screen was contacted via the electronic medical record and advised to obtain a high-resolution CT (HRCT) scan in the inspiratory prone position if clinically indicated. A follow-up message was sent after 30 days if no HRCT scan had been ordered. ILD status was assessed by using the multi-reader approach described earlier. ILD subtype was determined by multidisciplinary discussion following ILD clinic evaluation or chart review.

Test Performance Comparative Analysis

ILD prediction and test performance were compared between the ILD-Screen and common clinical and physiologic ILD features. Cough, dyspnea, and lung crackles were considered present when mentioned by a physician in the electronic medical record within 6 months of the PFT. Physiologic variables were dichotomized above and below the 80% predicted threshold.

ILD-Screen Clinical Effectiveness

ILD-Screen effectiveness was assessed by comparing prospectively identified patients with a positive ILD-Screen vs a historical control cohort of similarly identified patients undergoing PFT from 2015 to 2016. Metrics of effectiveness assessed in these cohorts were the proportion of patients with a positive ILD-Screen who underwent subsequent chest CT imaging and the time from a PFT to chest CT scan in those for whom chest CT imaging was performed. Follow-up time was censored at 24 months.

Statistical Analysis

Continuous variables are presented as mean $\pm$ SD. Categorical variables are presented as counts and percentages and were compared by using a χ^2 test. Logistic regression was used to assess predictors of ILD. Time to event comparison was performed by using a log-rank test and displayed graphically with the Kaplan-Meier estimator. Statistical significance was set at $P < .05$. All statistical analyses were performed by using Stata Release 16 (StataCorp).

Results

ILD-Screen Derivation

Of 1,271 patients undergoing paired PFT and chest CT imaging over the derivation period, 212 (16.7%) had evidence of ILD. Interobserver agreement for ILD presence on chest CT imaging was good ($\kappa = 0.79$). Compared vs those without ILD, case subjects were older, with a higher predominance of male subjects and lower TLC, FVC, and D_{LCO} (Table 1). ILD was observed in a higher number of patients referred for a pulmonary indication compared with a nonpulmonary indication.

Variables selected for the final ILD-Screen model were age, height, absolute TLC, FEV_1 and D_{LCO}, and indication risk category (IRC) (Table 2). No interaction variable, linear spline, or variable derivative was included in the final regression model, as none significantly improved model AUC. Based on final regression model beta-coefficients (e-Table 1), an ILD-Screen score was constructed by using the following equation:

$(0.0544 * \text{age}) - (0.7827 * \text{TLC (L)}) + (1.4567 * FEV_1 \text{(L)}) - (0.2336 * D_{LCO} \text{(mL/mm Hg/min)}) + (0.0694 * \text{height (cm)}) + (1.1645 * \text{IRC}) - 3.9$

With a model AUC of 0.91 (e-Fig 1A), each one-point increase in the ILD-Screen score increased ILD risk by nearly threefold (OR, 2.72; 95% CI, 2.4-3.09; $P < .001$). Test performance characteristics at various ILD-Screen score thresholds were assessed (e-Table 2), with a score ≥ 8 chosen to classify a positive ILD-Screen.

External Validation

Three hundred eighty-four patients at the CHUM met inclusion and exclusion criteria, including 68 (17.7%) with evidence of ILD on chest CT imaging. Compared vs those with ILD in the derivation cohort (Table 1), this cohort had a higher percentage of patients for whom white race was used as the reference group, along with better overall lung function. When applied in this cohort, the ILD-Screen showed a sensitivity and specificity of 0.76 and 0.82, respectively (Table 3). With a model AUC of 0.86 (e-Fig 1B), each one-point increase in the ILD-Screen score increased the odds of ILD by nearly twofold (OR, 1.89; 95% CI, 1.64-2.18; $P < .001$).

Prospective Validation

Of 2,002 patients completing a PFT during the screening period, 699 (34.9%) met inclusion criteria, including 132 of 699 (18.9%) with a positive ILD-Screen (Fig 1). Of ILD-Screen positive cases, 98 (74.2%) underwent HRCT imaging after this test was recommended to the PFT-ordering physician. Of 567 screen-negative cases, 245 (43.2%) underwent chest CT imaging. Incident ILD was identified in 18.4% (n = 63) of individuals who underwent chest CT scanning, with moderate interobserver agreement for the presence of ILD on chest CT scans in this cohort ($\kappa = 0.53$). ILD case subjects were older than those in the retrospective cohorts and had significantly better overall lung function (Table 1), with mean TLC and FVC in the normal range. Although a higher percentage of those with ILD had TLC, FVC, and D_{LCO} < 80% predicted, large minorities

of those without ILD also had these physiologic abnormalities.

Of prospectively identified ILD cases, 50 (79%) had a positive ILD-Screen. The ILD-Screen exhibited sensitivity and specificity of 0.79 and 0.83, respectively, when applied in this cohort (Table 3). With a model AUC of 0.87 (e-Fig 1C), each one-point increase in the ILD-Screen score increased the odds of ILD by nearly threefold (OR, 2.93; 95% CI, 2.18-3.95; $P < .001$).

When stratifying the cohort according to the presence of ILD symptoms (cough and/or dyspnea), test performance was relatively similar. Sensitivity and specificity of the ILD-Screen were 0.78 and 0.8, respectively, in the symptomatic group, and 0.86 and 0.89 in the asymptomatic group (e-Table 3). This analysis was limited by relatively few ILD cases (n = 7) in the asymptomatic group. When stratifying the cohort according to care setting (e-Table 4), variable ILD-Screen performance was observed. In the primary care setting, the ILD-Screen had a sensitivity and specificity of 0.81 and 0.8. In other care settings, the sensitivity and specificity were 0.73 and 0.86; there were relatively few ILD cases (n = 15) in non-primary care settings, however. The relative contribution of each variable comprising the ILD-Screen is shown in e-Table 5.

TABLE 1] Cohort PFT Characteristics

Characteristic	Derivation Cohort		External Validation Cohort		Prospective Validation Cohort	
	ILD Present (n = 212)	ILD Absent (n = 1,059)	ILD Present (n = 68)	ILD Absent (n = 316)	ILD Present (n = 63)	ILD Absent (n = 280)
Age	67.9 ± 12.1	60.5 ± 14.6	67.8 ± 11.4	62.4 ± 13.2	71.4 ± 9.7	61.3 ± 14.8
Male sex	107 (50.5)	481 (45.4)	47 (69.1)	162 (51.3)	30 (47.6)	124 (44.3)
Reference race						
White	182 (85.9)	870 (82.2)	66 (97.1)	302 (95.6)	56 (89.9)	236 (84.3)
African American	21 (9.9)	101 (9.5)	2 (2.9)	6 (1.9)	3 (4.8)	26 (9.3)
Hispanic	9 (4.3)	88 (8.3)	0	8 (2.5)	4 (6.3)	18 (6.4)
BMI, kg/m^2	28.6 ± 6.2	29.6 ± 7.3	29.2 ± 5.2	28 ± 6.4	30.4 ± 6.7	30.1 ± 7.7
Height, cm	167.5 ± 9.8	167.9 ± 10	164.5 ± 8.7	164.5 ± 9.1	166.6 ± 8.6	168 ± 10.3
Weight, kg	78.6 ± 19.2	81.9 ± 22.6	77.6 ± 15.2	74.2 ± 18.6	82.4 ± 21	83.8 ± 24
TLC, L	4.3 ± 1.3	5.4 ± 1.5	4.5 ± 1.3	5.5 ± 1.4	4.9 ± 1.3	5.4 ± 1.5
TLC (% predicted)	73.8 ± 19.2	93.4 ± 18	76.9 ± 18.6	97.5 ± 18	85.8 ± 16.2	91.1 ± 15.4
TLC < 80% predicted	130 (61.3	238 (22.5)	40 (58.8)	43 (13.6)	20 (31.8)	71 (25.4)
FEV$_1$, L	1.9 ± 0.7	2.2 ± 0.9	2.2 ± 0.7	2.2 ± 0.8	2 ± 0.6	2.2 ± 0.8
FEV$_1$ (% predicted)	74.2 ± 22	75.4 ± 22.3	85.3 ± 21.6	80.5 ± 23.2	81.6 ± 18.8	78.3 ± 20.1
FEV$_1$ < 80% predicted	127 ± 59.9	571 ± 53.9	26 ± 38.2	131 ± 41.5	28 ± 44.4	147 ± 52.5
FVC, L	2.6 ± 0.9	3.2 ± 1.1	2.7 ± 0.9	3.1 ± 1	2.9 ± 0.9	3.2 ± 1
FVC (% predicted)	75.2 ± 22.1	85.8 ± 20	77.7 ± 20.2	87.2 ± 19.5	95.9 ± 27.1	87.2 ± 19.5
FVC < 80% predicted	117 (55.2)	359 (33.9)	43 (63.2)	97 (30.7)	20 (31.8)	66 (23.6)
FEV$_1$/FVC	0.75 ± 0.1	0.68 ± 0.1	0.82 ± 0.1	0.7 ± 0.1	0.72 ± 0.1	0.7 ± 0.1
FEV$_1$/FVC > 0.8	76 (35.9)	175 (16.5)	42 (61.7)	86 (27.2)	15 (23.8)	40 (14.3)
D$_{LCO}$, mL/mm Hg/min	11.7 ± 4.4	17.3 ± 6	13.5 ± 5.2	16.9 ± 5.9	14.1 ± 4.9	18 ± 5.6
D$_{LCO}$ (% predicted)	49.5 ± 17.4	68.9 ± 18.3	58 ± 20	70.8 ± 19	61.4 ± 17.2	71.6 ± 15.8
D$_{LCO}$ < 80% predicted	201 (94.8)	769 (72.6)	57 (83.8)	217 (68.7)	55 (87.3)	198 (70.7)
PFT indication						
Pulmonary disease/ symptom	178 (84)	679 (64.1)	57 (83.8)	192 (60.8)	50 (79.4)	168 (60)
Nonpulmonary disease/ symptom	34 (16)	380 (35.9)	11 (16.2)	124 (39.2)	13 (20.6)	112 (40)

Values are expressed as mean ± SD or No. (%). D$_{LCO}$ = diffusing capacity of the lung for carbon monoxide; ILD = interstitial lung disease; PFT = pulmonary function test; TLC = total lung capacity.

TABLE 2] ILD Prevalence by PFT Indication

PFT Indication	Indication Risk Category[a]	Derivation Cohort (n = 1,271)	External Validation Cohort (n = 384)	Prospective Validation Cohort (n = 343)
Pulmonary disease				
Asthma	0	3/61 (4.9)	0/10	3/36 (8.3)
Bronchiectasis	1	6/40 (15)	0/1	0/0
COPD/emphysema	1	16/151 (10.6)	2/42 (4.8)	12/36 (33.3)
ILD	3	97/114 (85.1)	44/71 (62)	3/3 (100)
Other ILD features[b]	1	3/24 (12.5)	0/9	3/6 (50)
Nodule(s)	0	3/59 (5.1)	4/38 (10.5)	0/2
Pulmonary hypertension	2	7/24 (29.2)	5/50 (10)	3/18 (16.7)
Sarcoidosis	2	12/37 (32.4)	0/2	1/2 (50)
Other	0	3/75 (4)	2/12 (16.7)	1/2 (50)
Cough and/or dyspnea	1	28/272 (10.3)	0/0	24/113 (21.2)
Cardiac disease	0	5/79 (6.3)	0/0	2/18 (11.1)
Rheumatologic disease	3	8/17 (47.1)	2/15 (13.3)	6/10 (60)
Malignancy	0	16/247 (6.5)	8/108 (7.4)	5/79 (6.3)
Other	0	5/71 (7)	1/12 (8.3)	0/18

See Table 1 legend for expansion of abbreviation.
[a]Assigned based on derivation cohort prevalence.
[b]Restrictive lung disease, lung crackles, abnormal imaging, and hypoxemia.

Among incident ILD cases identified during the prospective screening period, the most common ILD subtypes were idiopathic pulmonary fibrosis (IPF), connective tissue disease-associated ILD (CTD-ILD), unclassifiable fibrotic ILD (U-ILD),[22] and unclassifiable nonfibrotic interstitial lung abnormalities[20,21] (ILAs) (Table 4). Fifty-four percent of patients (34 of 63) had evidence of pulmonary fibrosis, defined as the presence of honeycombing,

traction bronchiectasis, traction bronchiolectasis, or extensive reticulation affecting > 25% of the total lung field. The ILD-Screen accurately identified 100% (n = 34) of these patients. The ILD-Screen identified all patients with IPF and U-ILD, and a large majority of patients with CTD-ILD, other ILDs, and unclassifiable nonfibrotic ILAs. ILD-Screen performance was less accurate among those with smoking-related ILD.

TABLE 3] ILD-Screen Test Performance Across Cohorts

	Derivation		External Validation		Prospective Validation	
	ILD on HRCT Imaging					
ILD-Screen	Positive	Negative	Positive	Negative	Positive	Negative
Positive	182	198	52	58	50	48
Negative	30	861	16	258	13	232
Sensitivity	0.86		0.76		0.79	
Specificity	0.81		0.82		0.83	
PPV	0.48		0.47		0.51	
NPV	0.97		0.94		0.95	
LR, positive	4.6		4.9		4.6	
LR, negative	0.17		0.29		0.22	

ILD-RS = $(0.0544 * age) - (0.7827 * TLC (L)) + (1.4567 * FEV_1 (L)) - (0.2336 * D_{LCO} (mL/mm Hg/min)) + (0.0694 * height (cm)) + (1.1645 * IRC) - 3.9$. ILD-RS < 8 = negative; ILD-RS ≥ 8 = positive.

HRCT = high-resolution CT; LR = likelihood ratio; NPV = negative predictive value; PPV = positive predictive value; RS = risk score. See Table 1 legend for expansion of other abbreviation.
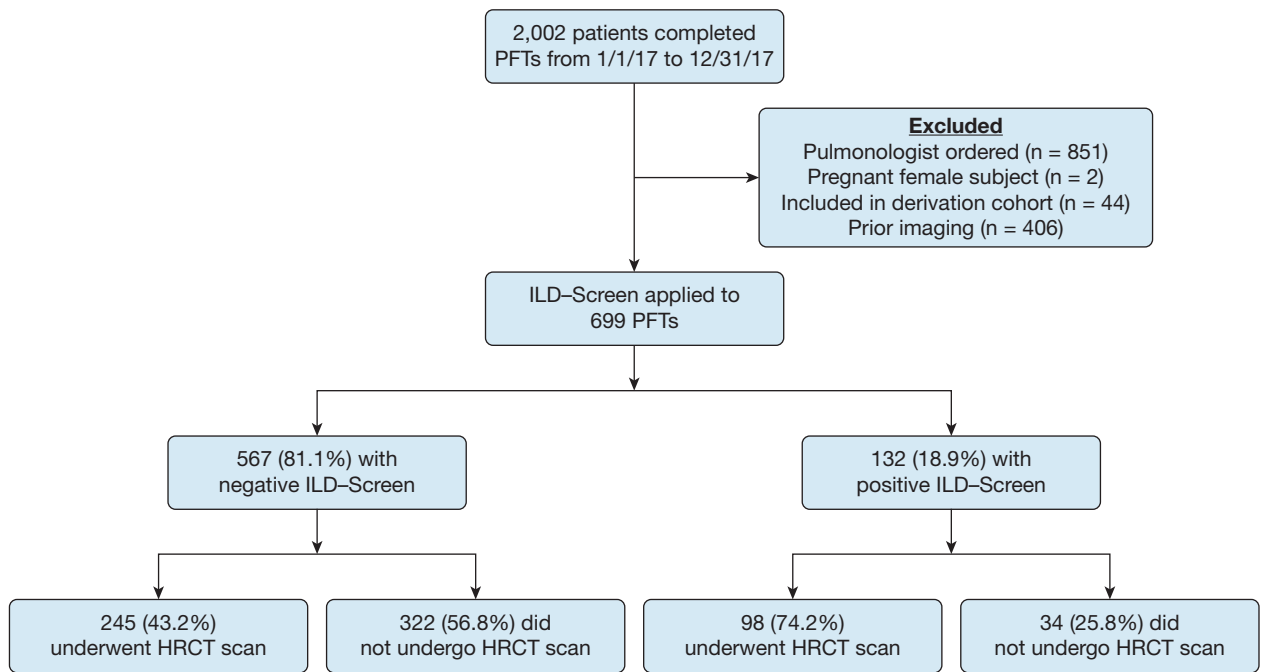
Figure 1 – *Consolidated Standards of Reporting Trials guidelines diagram for prospective ILD-Screen application. HRCT = high-resolution CT; ILD = interstitial lung disease; PFT = pulmonary function test.*

When assessing other radiologic findings in patients with a positive ILD-Screen who underwent chest CT imaging without evidence of ILD (n = 48), 15 patients (31%) had potentially actionable findings, including bulky adenopathy (n = 1), tree-in-bud opacities concerning for chronic infection (n = 3), features of chronic aspiration (n = 2), cylindrical bronchiectasis (n = 3), diffuse pulmonary nodules (n = 5), and acute pulmonary embolism (n = 1). Other findings that did not necessarily require a CT scan for diagnosis or characterization included centrilobular emphysema (n = 8), pulmonary edema (n = 2), pleural effusion (n = 1), acute pneumonia (n = 1), and post-pneumonectomy (n = 1). Twenty percent (20 of 98) of patients with a positive ILD-screen who underwent chest CT imaging had no radiologic abnormalities.

TABLE 4] Prospectively Applied ILD-Screen Performance Stratified According to ILD Subtype

Variable	ILD-Screen		% Correctly Classified
	Negative	Positive	
ILD classification			
Fibrotic ILD	0	34	100
Nonfibrotic ILD/ILAs	13	16	55
Diagnosis			
Idiopathic pulmonary fibrosis	0	6	100
Connective tissue disease-associated ILD	1	8	89
Unclassifiable fibrotic ILD	0	10	100
Unclassifiable nonfibrotic ILAs	5	13	72
Smoking-related ILD	4	5	56
Other ILD	3	8	73
Total	13	50	79

ILAs = interstitial lung abnormalities. See Table 1 legend for expansion of other abbreviation.

Test Performance Comparative Analysis

Among prospectively screened patients, all ILD features except dyspnea were associated with increased ILD risk (Table 5). Among clinical features, age > 60 years showed the strongest ILD association (OR, 5.7; 95% CI, 2.37-13.68) with a model AUC of 0.64. Among physiologic features, D$_{LCO}$ < 80% showed the strongest ILD association (OR, 2.85; 95% CI, 1.29-6.24), with a model AUC of 0.57. A positive ILD-Screen showed a stronger ILD association than any feature tested (OR, 18.6; 95% CI, 9.37-36.87), with a model AUC of 0.8. When comparing test performance characteristics among ILD features, only the ILD-Screen exhibited good sensitivity and specificity. Age > 60 years, cough, and D$_{LCO}$ < 80% demonstrated good sensitivity but poor specificity. Lung crackles and all other physiologic features exhibited good specificity but poor sensitivity.

ILD-Screen Clinical Effectiveness

Among ILD-Screen-positive patients, chest CT imaging was performed in 98 of 132 (74.2%) prospectively screened patients compared with 206 of 369 (55.8%) patients in the control cohort ($P = .003$). The median time to chest CT imaging was 5.6 months in the prospectively screened cohort compared with 21.1 months in the control cohort ($P < .001$) (Fig 2). A recommendation for HRCT imaging in those with a positive ILD-Screen was associated with a twofold increase in likelihood of chest CT images being obtained (OR, 2.19; 95% CI, 1.42-3.40; $P < .001$) compared with historical control subjects for whom this recommendation was not provided.

Discussion

In the current investigation, we developed the ILD-Screen, a diagnostic prediction tool for ILD derived from variables collected on PFTs. Prospective application of the ILD-Screen identified a large number of incident ILD cases, including all cases with pulmonary fibrosis and nearly all patients with IPF, CTD-ILD, and U-ILD. The ILD-Screen showed consistent test performance across two geographically diverse North American cohorts and

TABLE 5] Variables Associated With ILD in Prospective Validation Cohort

Variable	ILD (Positive) (n = 63)	ILD (Negative) (n = 280)	OR	95% CI	AUC	Sens	Spec	PPV	NPV	LR (Positive)	LR (Negative)
Clinical features											
Age > 60 y	57 (90.5)	175 (62.5)	5.7	2.37-13.68	0.64	0.9	0.38	0.25	0.95	1.45	0.26
Cough	52 (82.5)	172 (61.4)	2.96	1.48-5.94	0.61	0.82	0.39	0.23	0.91	1.34	0.46
Dyspnea	35 (55.6)	190 (67.9)	0.59	0.34-1.03	0.56	0.56	0.32	0.16	0.76	0.82	1.38
Crackles	15 (23.8)	30 (10.7)	2.6	1.30-5.20	0.57	0.24	0.89	0.33	0.84	2.18	0.85
Physiologic features											
TLC < 80% predicted	20 (31.8)	71 (25.4)	1.37	0.76-2.48	0.53	0.32	0.75	0.22	0.83	1.28	0.91
FVC < 80% predicted	20 (31.8)	66 (23.6)	1.58	0.83-2.74	0.55	0.32	0.76	0.23	0.83	1.33	0.89
D$_{LCO}$ < 80% predicted	55 (87.3)	198 (70.7)	2.85	1.29-6.24	0.57	0.87	0.29	0.22	0.91	1.23	0.45
TLC and FVC < 80% predicted	15 (23.8)	40 (14.3)	1.88	0.96-3.66	0.57	0.24	0.86	0.27	0.83	1.71	0.88
FVC and D$_{LCO}$ < 80% predicted	20 (31.8)	57 (20.4)	1.92	0.99-3.33	0.57	0.32	0.8	0.26	0.84	1.60	0.85
TLC and D$_{LCO}$ < 80% predicted	20 (31.8)	64 (22.9)	1.57	0.86-2.86	0.56	0.32	0.77	0.24	0.84	1.39	0.88
TLC and FVC and D$_{LCO}$ < 80% predicted	15 (23.8)	37 (13.2)	2.05	1.04-4.03	0.57	0.24	0.87	0.29	0.84	1.85	0.87
ILD-Screen											
Positive result	50 (79.4)	48 (17.1)	18.6	9.37-36.87	0.8	0.79	0.83	0.51	0.95	4.65	0.25

Values are expressed as No. (%). AUC = area under the curve. See Table 1 and 3 legends for expansion of other abbreviations.
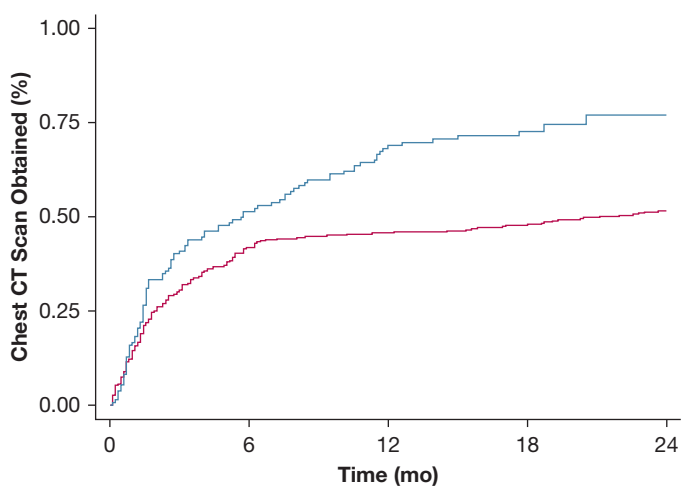
Figure 2 – *Time to chest CT scan in ILD-Screen-positive patients. A higher proportion of prospectively screened patients underwent chest CT imaging, and over a shorter time frame, compared with a historical control cohort (P < .001). See Figure 1 legend for expansion of abbreviation.*

when applied prospectively, thus suggesting good generalizability. This tool discriminated ILD from non-ILD conditions significantly better than any ILD clinical feature, including cough, dyspnea, and lung crackles, and categorically reduced % predicted TLC, FVC, and D$_{LCO}$. To our knowledge, the ILD-Screen represents the first validated diagnostic prediction tool for ILD.

Prospective application of the ILD-Screen identified 50 incident ILD cases at our institution over 1 year. It is unclear what percentage of patients would have undergone HRCT imaging without ILD-Screen application, but a 33% relative increase in chest CT imaging acquisition was observed in prospectively screened patients, and in a significantly shorter time frame, than those in a historical control cohort. These findings suggest that the ILD-Screen application has tangible benefits over the current standard of care by facilitating earlier HRCT assessment. We recently showed that substantial diagnostic delays occur in patients with ILD even after undergoing PFT showing features suggestive of ILD[12]; thus, the ILD-Screen represents a point of intervention to reduce the time from PFT to chest CT imaging. Although we relied on manual extraction of results of PFT and clinical data for this study, emerging software that efficiently extracts clinical data from the electronic medical record will be instrumental in systematizing the ILD-Screen and creating a perpetually increasing cohort to optimize its test performance.

When developing the ILD-Screen, we chose a balanced sensitivity and specificity threshold for a positive test result in an effort to capture a large majority of ILD cases while minimizing the number of patients who underwent unnecessary chest CT imaging. As expected, this approach resulted in modest likelihood ratios and a low positive predictive value given the low prevalence of ILD, with nearly one-half of ILD-Screen- positive patients having no evidence of ILD on HRCT imaging. Although this percentage of false-positive findings remains less than ideal, our data suggest that the decision to pursue chest CT imaging for any other common ILD feature, including cough, dyspnea, lung crackles, or restrictive physiology, would result in a significantly higher number of unnecessary chest CT scans.

Cough and dyspnea are among the most common clinical features in patients with ILD and should raise the suspicion of ILD when present.[23] Our data support these observations, as we found that > 90% of patients with ILD in our prospectively screened cohort endorsed one or both of these symptoms. However, the specificity of these symptoms for the presence of ILD was poor. Conversely, lung crackles are specific for the presence of ILD but showed poor sensitivity. Our data support the research of other investigators who have shown that lung crackles are strongly associated with pulmonary fibrosis.[24] Such findings have led some to propose lung auscultation as a screening tool for ILD.[25] Our data suggest, however, that this approach would miss a large

percentage of patients with ILD, as only 24% of patients had this finding documented by a physician prior to PFT. Interestingly, we found that age > 60 years had 90% sensitivity for detecting ILD, albeit with poor specificity. This finding suggests that ILD-Screen optimization within categorical age groups may further improve test performance.

Physiologic features of ILD, including categorically reduced TLC, FVC, and DLCO, also showed relatively poor test performance for identifying ILD. We found that a DLCO < 80% predicted was highly sensitive for ILD but poorly specific. All combinations of categorical PFT abnormalities assessed, including the ILD "phenotype" of low TLC, FVC, and DLCO,[16] showed good specificity but poor sensitivity. These data are in line with others who have shown that categorical reductions in TLC, FVC, and DLCO have poor overall test performance for discriminating ILD in patients with scleroderma.[26] The ILD-Screen outperformed all of these metrics and suggests that a focus on any single or combined PFT abnormality fails to capture important physiologic interplay that is accounted for by the ILD-Screen model. Taking into account PFT indication also improved ILD prediction. Although those with an ILD indication for PFT had an understandably high ILD risk, other indications displayed similarly differential ILD risk. For example, those with a rheumatologic indication had strikingly high ILD risk, whereas those undergoing PFT for asthma or malignancy displayed markedly lower risk. This prior knowledge improves ILD risk estimation and with larger datasets, specific point estimates for common indications can be ascertained, thus further improving subsequent ILD-Screen iterations.

Within ILD subtypes, we found that the ILD-Screen was most effective at identifying fibrotic ILD, as this tool correctly classified all patient evidence of pulmonary fibrosis, including the majority of patients with IPF, CTD-ILD, and U-ILD, which are among the most common ILDs[27] and have high morbidity and mortality.[22,23,28] The ILD-Screen was less reliable in detecting nonfibrotic ILDs, including unclassifiable nonfibrotic ILAs. We chose to define ILD based on research by other investigators who had first characterized ILAs.[20,21] These findings seem to be the earliest radiographic finding in ILD and are linked to an increased risk of death,[29] hospitalization,[30] and disease progression.[19] The ILD-Screen correctly classified nearly 60% of patients with ILAs, who by definition had at least

mild decrements in lung function, suggesting that this tool may be useful in identifying clinically relevant ILAs. Test performance was poorer in patients with smoking-related ILD, which may reflect a lower FEV_1 in these patients due to concurrent airflow obstruction.

The current study had several limitations. First, the cohorts assessed in this study were likely enriched for ILD cases, introducing selection bias. A large majority of prospectively screened patients endorsed cough and/or dyspnea, and the prevalence of ILD in this population was more than twice that described in the general population.[19] In addition, a large number of patients in the retrospectively screened cohorts had prevalent ILD. However, by excluding prevalent ILD cases from our prospectively screened cohort, the current data support the effectiveness of this tool in identifying incident ILD cases in nonpulmonology settings. Additional research is needed to determine the utility of ILD-Screen application in asymptomatic and average risk populations. Another limitation stems from the reliance on retrospective chart review to ascertain the status of ILD signs and symptoms. Although most patients had extensive medical records available for review, some patients may have experienced ILD signs and symptoms without documentation by a physician. In addition, some patients with a positive ILD-Screen did not undergo chest CT imaging. It is possible that ILD prevalence was lower in this group, which would alter ILD-Screen test performance. Finally, we found that ILD-Screen test performance varied across care settings, with the best overall performance in the primary care setting, with a worse sensitivity but better specificity in other care settings. This finding suggests that that the ILD-Screen may best be used in these populations and that discipline-specific modifications to the ILD-Screen may be warranted once larger cohorts are assembled.

Conclusions

The current study showed that the ILD-Screen is a more reliable predictor of ILD than hallmark clinical and physiologic features of ILD among patients undergoing PFT for diverse indications and across various care settings. This tool represents a promising modality to begin identifying at-risk patients who should be referred for HRCT imaging, as such an intervention has the potential to reduce diagnostic delays, facilitate earlier intervention, and improve patient outcomes.

References

1. Fischer A, Kong AM, Swigris JJ, Cole AL, Raimundo K. All-cause healthcare costs and mortality in patients with systemic sclerosis with lung involvement. *J Rheumatol.* 2018;45(2):235-241.

2. Collard HR, Chen SY, Yeh WS, et al. Health care utilization and costs of idiopathic pulmonary fibrosis in U.S. Medicare beneficiaries aged 65 years and older. *Ann Am Thorac Soc.* 2015;12(7):981-987.

3. Best AC, Meng J, Lynch AM, et al. Idiopathic pulmonary fibrosis: physiologic tests, quantitative CT indexes, and CT visual scores as predictors of mortality. *Radiology.* 2008;246(3):935-940.

4. Mooney JJ, Elicker BM, Urbania TH, et al. Radiographic fibrosis score predicts survival in hypersensitivity pneumonitis. *Chest.* 2013;144(2):586-592.

5. Fischer A, Brown KK, Du Bois RM, et al. Mycophenolate mofetil improves lung function in connective tissue disease-associated interstitial lung disease. *J Rheumatol.* 2013;40(5):640-646.

6. Lederer DJ, Martinez FJ. Idiopathic pulmonary fibrosis. *N Engl J Med.* 2018;379(8):797-798.

7. Morisset J, Johannson KA, Vittinghoff E, et al. Use of mycophenolate mofetil or azathioprine for the management of chronic hypersensitivity pneumonitis. *Chest.* 2017;151(3):619-625.

8. Distler O, Highland KB, Gahlemann M, et al. Nintedanib for systemic sclerosis-associated interstitial lung disease. *N Engl J Med.* 2019;380(26):2518-2528.

9. Cosgrove GP, Bianchi P, Danese S, Lederer DJ. Barriers to timely diagnosis of interstitial lung disease in the real world: the INTENSITY survey. *BMC Pulmon Med.* 2018;18(1):9.

10. Collard HR, Tino G, Noble PW, et al. Patient experiences with pulmonary fibrosis. *Respir Med.* 2007;101(6):1350-1354.

11. Hoyer N, Prior TS, Bendstrup E, Wilcke T, Shaker SB. Risk factors for diagnostic delay in idiopathic pulmonary fibrosis. *Respir Res.* 2019;20(1):103.

12. Pritchard D, Adegunsoye A, Lafond E, et al. Diagnostic test interpretation and referral delay in patients with interstitial lung disease. *Respir Res.* 2019;20(1):253.

13. Ryerson CJ, Vittinghoff E, Ley B, et al. Predicting survival across chronic interstitial lung disease: the ILD-GAP model. *Chest.* 2014;145(4):723-728.

14. Ley B, Ryerson CJ, Vittinghoff E, et al. A multidimensional index and staging system for idiopathic pulmonary fibrosis. *Ann Intern Med.* 2012;156(10):684-691.

15. Wells AU, Desai SR, Rubens MB, et al. Idiopathic pulmonary fibrosis: a composite physiologic index derived from disease extent observed by computed tomography. *Am J Respir Crit Care Med.* 2003;167(7):962-969.

16. Martinez FJ, Flaherty K. Pulmonary function testing in idiopathic interstitial pneumonias. *Proceed Am Thorac Soc.* 2006;3(4):315-321.

17. Kadoch M, Kitich A, Alqalyoobi S, et al. Interstitial lung abnormality is prevalent and associated with worse outcome in patients undergoing transcatheter aortic valve replacement. *Respir Med.* 2018;137:55-60.

18. Oldham JM, Adegunsoye A, Khera S, et al. Underreporting of interstitial lung abnormalities on lung cancer screening computed tomography. *Ann Am Thorac Soc.* 2018;15(6):764-766.

19. Araki T, Putman RK, Hatabu H, et al. Development and progression of interstitial lung abnormalities in the Framingham Heart Study. *Am J Respir Crit Care Med.* 2016;194(12):1514-1522.

20. Hunninghake GM, Hatabu H, Okajima Y, et al. MUC5B promoter polymorphism and interstitial lung abnormalities. *N Engl J Med.* 2013;368(23):2192-2200.

21. Washko GR, Hunninghake GM, Fernandez IE, et al. Lung volumes and emphysema in smokers with interstitial lung abnormalities. *N Engl J Med.* 2011;364(10):897-906.

22. Ryerson CJ, Urbania TH, Richeldi L, et al. Prevalence and prognosis of unclassifiable interstitial lung disease. *Eur Respir J.* 2013;42(3):750-757.

23. Raghu G, Remy-Jardin M, Myers JL, et al. Diagnosis of idiopathic pulmonary fibrosis. An official ATS/ERS/JRS/ALAT clinical practice guideline. *Am J Respir Crit Care Med.* 2018;198(5):e44-e68.

24. Sgalla G, Walsh SLF, Sverzellati N, et al. Velcro-type" crackles predict specific radiologic features of fibrotic interstitial lung disease. *BMC Pulm Med.* 2018;18(1):103.

25. Cottin V, Cordier JF. Velcro crackles: the key for early diagnosis of idiopathic pulmonary fibrosis? *The European respiratory journal.* 2012;40(3):519-521.

26. Suliman YA, Dobrota R, Huscher D, et al. Brief report: pulmonary function tests: high rate of false-negative results in the early detection and screening of scleroderma-related interstitial lung disease. *Arthritis Rheumatol.* 2015;67(12):3256-3261.

27. Duchemann B, Annesi-Maesano I, Jacobe de Naurois C, et al. Prevalence and incidence of interstitial lung diseases in a multi-ethnic county of Greater Paris. *Eur Respir J.* 2017;50(2).

28. Strand MJ, Sprunger D, Cosgrove GP, et al. Pulmonary function and survival in idiopathic vs secondary usual interstitial pneumonia. *Chest.* 2014;146(3):775-785.

29. Putman RK, Hatabu H, Araki T, et al. Association between interstitial lung abnormalities and all-cause mortality. *JAMA.* 2016;315(7):672-681.

30. Podolanczuk AJ, Oelsner EC, Barr RG, et al. High-attenuation areas on chest computed tomography and clinical respiratory outcomes in community-dwelling adults. *Am J Respir Crit Care Med.* 2017;196(11):1434-1442.