

Human epididymis protein 4 is a new biomarker to predict the prognosis of progressive fibrosing interstitial lung disease

Naoki Nishiyama

Tokyo Ika Shika Daigaku - Yushima Campus

Masahiro Masuo

Tokyo Ika Shika Daigaku - Yushima Campus

Yoshihisa Nukui

Tokyo Ika Shika Daigaku - Yushima Campus

Tomoya Tateishi

Tokyo Ika Shika Daigaku - Yushima Campus

Mitsuhiro Kishino

Tokyo Ika Shika Daigaku

Ukhide Tateishi

Tokyo Ika Shika Daigaku - Yushima Campus

Kaori Morota

Abbott Laboratories Limited

Kazuyuki Ohbo

Yokohama Shiritsu Daigaku

Yasunari Miyazaki (✉ miyazaki.pilm@tmd.ac.jp)

Tokyo Ika Shika Daigaku - Yushima Campus <https://orcid.org/0000-0002-9073-9815>

Research article

Keywords: human epididymis protein 4, progressive fibrosing interstitial lung disease, honeycombing, prognosis, biomarker

Posted Date: May 18th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-18084/v2>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at Respiratory Investigation on January 1st, 2021. See the published version at <https://doi.org/10.1016/j.resinv.2020.08.002>.

Abstract

BACKGROUND The clinical course and prognosis of progressive fibrosing interstitial lung diseases (PF-ILDs) vary from individual to individual. Predictive serum biomarkers for the management of the disease are needed. Serum human epididymis protein 4 (HE4) has been reported to be elevated in patients with IPF, yet its clinical utility has not been elucidated. We evaluated whether serum HE4 could be a biomarker for patients with PF-ILD.

METHODS Serum HE4 was measured in a retrospective study that consisted of 34 patients with PF-ILD and 40 healthy volunteers. The relationship between serum HE4 levels and clinical parameters or prognosis was investigated. To validate the significance of the results, a prospective observational study that consisted of 37 patients with PF-ILD and 40 control patients without PF-ILD was performed.

RESULTS Serum HE4 was higher in patients with PF-ILD than in health volunteers ($P < 0.01$). A correlation of the serum HE4 levels with the extent of honeycombing on chest high-resolution computed tomography was identified ($r = 0.41$, $P = 0.015$). In multivariate analysis by the Cox proportional hazard model, higher HE4 levels (> 238 pmol/l) were associated with elevated mortality risk (HR 7.27, 95% CI 1.56-34.0, $P = 0.01$ in the derivation cohort; HR 44.3, 95% CI 4.19-468, $P < 0.01$ in validation cohort).

CONCLUSIONS Serum HE4 levels may serve as a new diagnostic and prognostic biomarker for patients with PF-ILD.

Introduction

Interstitial lung diseases (ILDs) encompass a large variety of diseases that are caused by a known etiology (such as autoimmune disease, vasculitis, drugs, tumors and occupational or environmental exposure) or unknown etiology (idiopathic). Among ILDs, idiopathic pulmonary fibrosis (IPF), which is a representative chronic fibrosing lung disease, has a poor prognosis with a median survival of 2-3 years [1]. Some types of ILDs other than IPF (such as chronic hypersensitivity pneumonitis (CHP), connective tissue disease-associated ILD (CTD-ILD), idiopathic nonspecific interstitial pneumonia (iNSIP), unclassifiable ILD and asbestosis) are also known to have progressive disease behaviors similar to IPF [2, 3]. Although ILDs with a progressive phenotype are termed progressive fibrosing-ILD (PF-ILD), these disease behaviors are unpredictable and widely variable. Therefore, clinical biomarkers to predict progression are required.

Human epididymis protein 4 (HE4), also known as whey-acidic-protein (WAP) four-disulfide core domain protein 2 (WFDC2) containing two WAP domains, was first identified as a secretory protein in the human epididymis [4, 5]. Several studies have revealed that HE4 is expressed in multiple tissues, such as the respiratory tract, kidney, prostate, and ovarian cancer [6-9]. Although its biological role remains largely unknown, it has been suggested that HE4 has a functional role in the formation of fibrosis. The correlation between renal fibrosis and serum HE4 levels in patients with kidney disease and the function of HE4 in inhibiting matrix metalloproteinase 2 and 9 (MMP2, MMP9), inducing the degradation of type I

collagen, have been reported [10]. For the lung, serum HE4 levels have been reported to be higher in patients with IPF than in patients without respiratory disease [11]. However, the clinical association between serum HE4 levels and pulmonary fibrosis has not been elucidated.

Therefore, we performed this study to evaluate whether serum levels of HE4 are elevated in patients with PF-ILD or can be used as a biomarker. We retrospectively measured the serum levels of HE4 and analyzed survival data, high-resolution computed tomography (HRCT) findings, and other clinical variables in correlation with HE4. The usefulness of HE4 as a biomarker was confirmed in a prospective cohort study.

Methods

Study Design and Population

This study was approved by the Institutional Review Board at the Tokyo Medical and Dental University (TMDU) in accordance with the Declaration of Helsinki (approval number M2019-010). The study consisted of two cohorts. First, a retrospective medical record-based observational cohort study, which served as a derivation cohort, was performed. One hundred two patients with ILD who were admitted to TMDU between June 2010 and July 2011 were included. Among the 102 patients, serum samples from 47 patients were stored. A total of 34 patients who were diagnosed with PF-ILD were analyzed (Figure 1A). Forty staff members from our hospital and department were evaluated as healthy volunteers (HVs). Second, a prospective observational cohort study, which served as a validation cohort, was performed. Informed consent from all of the patients was obtained. In addition to the patients recruited at the time of diagnosis, patients attending our outpatient department on a routine basis were recruited from December 2012 to July 2014. Patients with bronchial asthma, chronic obstructive pulmonary disease, and bronchiectasis were also recruited as disease controls during the same study period. Serum samples were collected from 39 patients with PF-ILD and 40 control subjects during the recruitment period (Figure 1B). Patients with cancer at baseline and an estimated glomerular filtration rate (eGFR) of less than 50 ml/min/1.73m² were excluded in the two cohorts. None of the patients experienced any episode of acute exacerbation of ILD prior to recruitment or overlapped in the two cohorts.

Clinical data of patients were included age, gender, smoking history, laboratory findings, physiological test results, radiological findings, and the etiology of ILD. The outcome of patients with PF-ILD was followed from baseline to 3 years. The clinical outcome assessed was all-cause death.

The diagnosis for IPF and non-IPF IIPs (NSIP and unclassified) was based on the criteria from the American Thoracic Society and European Respiratory Society [12, 13]. CTD was diagnosed based on the diagnostic criteria commonly used for each disease [14-16]. Chronic HP was diagnosed based on combinations of clinical, radiological, and histological criteria [17, 18]. Acute exacerbation of ILD was diagnosed based on diagnostic criteria for acute exacerbation of IPF [19].

Blood samples for measuring HE4 were taken using a routine procedure, centrifuged, aliquoted, and stored at -20°C or below until analysis. The concentration of serum HE4 was measured by chemiluminescent microparticle immunoassay on an ARCHITECT® i2000 analyzer (Abbott Japan LLC).

Evaluation of high-resolution computed tomography

HRCT scans were analyzed at the level of the aortic arch, carina trachea, and right pulmonary vein in each lung. The following 7 categories that were classified by parenchymal abnormalities were evaluated by one expert chest radiologist (M.K.) and one pulmonary specialist (Y.N.) with no knowledge of the clinical information: reticulation, honeycombing, centrilobular nodules, ground-glass opacity, consolidation, emphysema, and traction bronchiectasis (TBE). The extent of each radiologic abnormality, excluding TBE, was quantified as a percentage of lung parenchyma affected, to the nearest 5%, in each of the six lung zones. TBE was graded on a scale of 0-3 as follows: 0 = none; 1 = mild; 2 = moderate; and 3 = severe (Supplementary Figure S1) [20]. The average total score for each variable was calculated as the mean score of the six zones. The readings of the two observers on the extent of each abnormality were combined by calculating the averages. Interobserver agreement on the extent of each abnormality was evaluated with the Spearman rank correlation coefficient.

Evaluation of histological findings

Lung tissues were obtained from 8 patients with PF-ILD who underwent SLB from 2005 to 2013 in our institution. The lung tissues were retrospectively selected and analyzed separately from this observational study. Control lung tissues were obtained from normal parts of lungs excised in lung cancer surgeries.

The lung tissue sections were subjected to antigen retrieval by autoclaving for 20 minutes in citrate buffer after deparaffinization. Endogenous peroxidase was quenched with 3% H₂O₂, and 5% goat serum and 1% BSA were used for blocking. After the addition of rabbit monoclonal antibody to HE4 (Abcam, Japan, Cat. No. ab200828) overnight at 4°C, the sections were incubated with a biotinylated secondary antibody (Vector Laboratories) for 60 minutes followed by ABC reagent (VECTASATAIN) for 30 minutes. Next, the sections were incubated in 3,3'-diaminobenzidine tetrahydrochloride (DAB; VECTOR) solution and counterstained with Meyers' hematoxylin solution. Negative controls were incubated without primary antibody.

Statistical analysis

Differences between the groups were analyzed by Student's t-test, Welch's t test, Mann-Whitney U test, or Pearson's chi-square test. Correlations were determined by the Spearman rank test. The most suitable cutoff level for serum HE4 levels was determined using a receiver operating characteristic (ROC) curve analysis with the highest Youden index (i.e., sensitivity × specificity - 1). The Kaplan-Meier method was used to assess survival curves with GraphPad Prism version 7 (Graph Pad Software Inc., La Jolla, CA, USA). The log-rank test was used to evaluate the statistical significance of differences between the higher HE4 and lower HE4 groups. Cox proportional hazards models were used to examine the influences of serum HE4 levels on the prognosis of patients with PF-ILD. Statistical analyses were performed using EZR software version 1.37 (Saitama Medical Center, Jichi Medical University, Saitama, Japan) [21]. A *P* value < 0.05 was considered statistically significant.

Results

Patient characteristics

The patient characteristics in the two cohorts are shown in Table 1. Patients with PF-ILD were significantly older than healthy volunteers (*P* < 0.01) in the derivation cohort. A greater proportion of patients with PF-ILD in the two cohorts were male (55.8% in the derivation cohort and 64.1% in the validation cohort). In the validation cohort, patients with PF-ILD had a significantly lower smoking history than control patients. The control subjects in the validation cohort consisted of 11 patients with bronchial asthma, 15 with chronic obstructive pulmonary disease (COPD), and 14 with bronchiectasis. The etiology of PF-ILDs mainly consisted of CHP (16/34, 47%) in the derivation cohort and IIPs (28/39, 71.7%) in the validation cohort. Krebs von den Lugen-6 (KL-6) and the estimated glomerular filtration rate (eGFR) were significantly higher in patients with PF-ILD in the derivation cohort than in those in the validation cohort (KL-6, *P* = 0.035; SP-D, *P* = 0.011). Diffusing capacity divided by the alveolar volume (DLCO/VA) was lower in the derivation cohort than that in the validation cohort (*P* = 0.002). Eighteen patients (steroid, *n* = 17; tacrolimus, *n* = 1) were undergoing immunosuppressive treatment at baseline in the derivation cohort. None of the patients had taken any treatment for PF-ILD at baseline in the validation cohort.

Correlation between HE4 and clinical variables

The serum HE4 levels were significantly higher in the patients with PF-ILD than in the control subjects in both cohorts (PF-ILD in the derivation cohort vs. HV, 180.8 ± 72.9 vs. 44.1 ± 17.2 pmol/l, *P* < 0.001; PF-ILD in the validation cohort vs. control patients, 140.9 ± 56.3 vs. 70.5 ± 29.4 pmol/l, *P* < 0.001) (Figure 2).

In the derivation cohort, serum HE4 levels in patients with PF-ILD were significantly correlated with age (*r* = 0.49, *P* < 0.01) and with the extent of honeycombing on HRCT (*r* = 0.41, *P* = 0.015) but not with serum KL-6 or SP-D levels (Table 2). The assessments of the various abnormalities on HRCT showed good to excellent interobserver agreement (*r* = 0.73-0.81, *P* < 0.01, Supplementary Table S1).

Identification of prognostic biomarkers

During the follow-up period, 9 patients (9/34, 26.4%) in the derivation cohort died. The causes of death were acute exacerbation (n = 5), chronic respiratory failure (n = 3), and malignant neoplasm (n = 1). First, to identify factors associated with the occurrence of acute exacerbation, characteristics were compared between patients who developed acute exacerbation and those who did not, but no factors related to acute exacerbation were identified (data not shown). Second, to identify factors associated with survival, characteristics were compared between survivors and nonsurvivors (Table 3). A serum HE4 level was the only factor significantly associated with the prognosis of PF-ILD. ROC analysis was conducted to determine the best cutoff value of serum HE4 level between the survivors and nonsurvivors (cutoff 238 pmol/l, AUC 0.76 (95% confidence interval (CI): 0.576 - 0.944)). The patients were divided into a higher HE4 group (n = 9, HE4 > 238 pmol/l) and a lower HE4 group (n = 25, HE4 ≤ 238 pmol/l) to analyze the survival using the Kaplan-Meier method. The log-rank test showed a significant difference in survival between the two groups ($P < 0.01$) (Figure 3A). The hazard ratio (HR) for mortality was calculated using the Cox proportional hazards model (Table 4). An association with an increased risk of mortality was observed in patients with serum HE4 levels above 238 pmol/l in the univariate model (HR 6.54, 95% CI: 1.71-25.0, $P < 0.01$), while no association with mortality risk was found in KL-6 and SP-D levels. Variables with a P value less than or equal to 0.20 in the univariate analysis were included in the multivariate analysis. Older age (≥ 65 , HR 12.0, 95% CI: 1.31-109, $P = 0.027$) and higher serum HE4 levels (> 238 pmol/l, HR 7.27, 95% CI: 1.56-34.0, $P = 0.012$) showed statistical significance in the multivariate analysis.

Furthermore, the cutoff value of serum HE4 was validated for the prediction of survival in the validation cohort. Ten patients (10/39, 25.6%) died from recruitment to 3 years. The causes of death were acute exacerbation (n = 7), and chronic respiratory failure (n = 3). The log-rank test analyzed by the Kaplan-Meier method for survival between the higher (n = 4, HE4 > 238 pmol/l) and lower HE4 (n = 35, HE4 ≤ 238 pmol/l) groups showed a significant difference between the two groups ($P < 0.01$) (Figure 3B). Multivariate analysis by the Cox proportional hazard model for mortality, using the same variables included in the analysis of the derivation cohort, identified higher HE4 as an independent predictor for mortality (HR 44.3, 95% CI: 4.19-468, $P < 0.01$).

Immunohistochemical analysis

The characteristics of the patients whose lung tissues were obtained are shown in e-Table 2. We investigated the expression of HE4 in the lung tissues obtained from 1 patient with IPF, 6 patients with IIPs, 3 with CHP, and 1 with CTD-IP. Normal lung tissues were obtained from 2 patients who underwent surgical resection for lung cancer. Representative HE4 expression levels in the lung tissues obtained from a representative control subject, a patient with IPF with a UIP pattern (48-year-old male), and a patient with IIP with an NSIP pattern (50-year-old male) are shown in Figure 4. In the normal lung tissue, HE4 was not

observed (Figure 4A, 4D). In the lung tissue of the patient with IPF, HE4 was expressed in metaplastic bronchiolized epithelium lining regions of honeycombing and bronchiolar epithelium (Figure 4B, 4E). In the lung tissue of the patient with NSIP, HE4 was expressed only in bronchiolar epithelium (Figure 4C, 4F). HE4 was not detected in the alveolar epithelium or fibroblast in any lung tissues of patients with PF-ILD.

Discussion

Investigators who are researching PF-ILD have suggested that we should figure out biomarkers that can be easily obtained through simple, noninvasive examinations. In this study, we showed that serum HE4 levels were higher in patients with PF-ILD than in control subjects and that higher HE4 levels were associated with worsened prognosis of patients with PF-ILD. This is the first report to indicate that HE4 is valuable as a diagnostic and prognostic biomarker of PF-ILD.

PF-ILD is a progressive fibrosing phenotype with poor prognosis, but the individual disease course is heterogeneous [3]. The duration from the time when the progressive fibrosing phenotype was detected by physicians to patient death was reported as 30-45 months [22]. Therefore, biomarkers to predict early disease progression or death could be useful for making decisions during treatment. In the clinical setting, the methods of assessing progression are pulmonary function tests, laboratory data, evaluation of symptoms, the 6-min walk test and chest imaging [23]. Lower forced vital capacity (FVC) and a decline in FVC > 10% predicted are established predictors of mortality in patients with IPF [24]. Among several serum biomarkers, higher KL-6 levels (>1000 U/ml) are associated with increased mortality in patients with IPF [25, 26]. However, KL-6 was not identified as a prognostic biomarker in the present study. Since patients with CHP who have been reported to have seasonal variation in the elevation of KL-6 levels [27] were predominantly included in our study, the difference in the etiology of ILD might have led to the result.

Elevation of serum HE4 levels has been reported in patients with lung cancer [28], ovarian cancer [29], and chronic kidney disease [30], and has already been applied to the diagnosis of ovarian cancer in clinical practice. Therefore, these patients were excluded from this study to avoid confounding factors. Another previous study has showed a positive correlation between serum HE4 levels and age [31], which was similar to our result. In our derivation cohort, the serum HE4 levels were significantly higher in the patients with PF-ILD than in the healthy volunteers, but a significant difference in age between the two groups might affect this result. Therefore, in our validation cohort, control subjects who were close to the age of patients with PF-ILD were enrolled. Although the proportion of subjects with past smoking history was higher in control subjects than in patients with PF-ILD, a positive correlation between past smoking history and serum HE4 levels was identified [31]. Taking into account these factors, patients with PF-ILD had significantly higher serum HE4 levels than control subjects .

The reason for the association between HE4 levels and prognosis is currently unclear. The expression of HE4 in the lung has been identified in epithelial cells of the airway but not in peripheral lung [7]. In our immunohistochemistry analysis of lung specimens, HE4 was not observed in the bronchiolar epithelium or lung parenchyma of normal lungs. However, we found that HE4 was expressed in

metaplastic bronchiolized epithelium lining regions of honeycombing and/or bronchiolar epithelium in PF-ILD lungs. Since HE4 is a secretory protein [32], this result might be associated with elevation of serum HE4 concentration in patients with PF-ILD. HE4 suppresses the activity of multiple proteases and especially inhibits MMP2/9 mediated degradation of type I collagen in fibrotic kidneys [10]. MMP2/9 have also been shown to be upregulated in IPF [33], and HE4 might be associated with lung fibrosis by the a mechanism similar to the profibrotic process identified in the fibrotic kidney. Although the role of HE4 in the formation of pulmonary fibrosis does not go beyond the range of speculation, we found a correlation between serum HE4 levels and the extent of honeycombing on a HRCT findings.

Honeycombing is one of the pivotal CT characteristics for the diagnosis of IPF and is defined as clustered cystic airspaces of the same diameters (approximately 3–10 mm) with thick, well-defined walls in the subpleural area [34]. The clinical significance has been suggested that pathological honeycombing reflects the end-stage of fibrotic changes [35]. The correlation between the serum HE4 levels and the HRCT findings implies that HE4 may play an important role in ongoing fibrosis and disease progression in PF-ILD.

The present study had several limitations. First, this study was a single-center study with a small number of analyses. Second, several types of PF-ILD were examined altogether in our analysis. We did not examine potential differences in the elevation of HE4 according to the etiology of PF-ILD. Third, our longitudinal examination of patients with PF-ILD for HE4 was insufficient. Repeated measurements of serum HE4 levels might have added very useful data. Fifth, there was a difference in the characteristics of control subjects between the derivation cohort and the validation cohort. Further studies to clarify the significance of the biomarker and the role of HE4 in pulmonary fibrosis are needed.

Conclusion

HE4 has a predictive significance for prognosis in PF-ILD. Measuring HE4 may be a supporting tool for pulmonologists to diagnose PF-ILD and predict the disease course of PF-ILD.

Abbreviation List

CHP = chronic hypersensitivity pneumonia; CTD = connective tissue disease; DLCO = diffusing capacity for carbon monoxide; eGFR = estimated glomerular filtration rate; FEV1 = forced expiratory volume; FVC = forced vital capacity; HE4 = human epididymis protein 4; HV = healthy volunteer; HRCT = high-resolution computed tomography; IIPs = idiopathic interstitial pneumonias; ILD = interstitial lung disease; IPF = idiopathic pulmonary fibrosis; KL-6 = Krebs von den Lungen 6; NSIP = nonspecific interstitial pneumonia; PF-ILD = progressive fibrosing- interstitial lung disease; SP-D = surfactant protein D; UIP = usual interstitial pneumonia; VA = alveolar volume

Declarations

Ethics approval and consent to participate

This study was approved by the ethics committee of Tokyo Dental and Medical University (approval number M2019-010). Informed consent from subjects in the retrospective cohort was obtained in the form of opt-out. Written informed consent was obtained from subjects in the prospective cohort before recruitment.

Consent for publication

Not applicable.

Availability of data and materials

All the data are available from the corresponding author on reasonable request.

Competing interests

Abbott Japan LLC covered the cost for serum HE4 measurement.

The authors declare that they have no competing interests.

Funding

This work was supported by JSPS KAKENHI [Grant Number 16K09576].

Authors' contributions

Conceptualization: N.N., Y.M. Data curation: N.N., M.M. Formal analysis: N.N., Y.N, M.K.,U.T., K.M. Funding acquisition: Y.M. Investigation: N.N., M.M., Y.N, M.K.,U.T., K.M., K.O. Methodology: M.M. Project administration: K.M. Y.M. Supervision: Y.N., T.T. Validation: N.N. Writing – original draft: N.N. Writing – review & editing: T.T., K.O.

Acknowledgements

Not applicable.

References

1. Kim DS, Collard HR, King TE, Jr.: **Classification and natural history of the idiopathic interstitial pneumonias.** *Proc Am Thorac Soc* 2006, **3**(4):285-292.
2. Cottin V, Hirani NA, Hotchkiss DL, Nambiar AM, Ogura T, Otaola M, Skowasch D, Park JS, Poonyagariyagorn HK, Wuyts W *et al.*: **Presentation, diagnosis and clinical course of the spectrum of progressive-fibrosing interstitial lung diseases.** *European respiratory review : an official journal of the European Respiratory Society* 2018, **27**(150).
3. Kolb M, Vasakova M: **The natural history of progressive fibrosing interstitial lung diseases.** *Respiratory research* 2019, **20**(1):57.
4. Kirchhoff C, Habben I, Ivell R, Krull N: **A major human epididymis-specific cDNA encodes a protein with sequence homology to extracellular proteinase inhibitors.** *Biology of reproduction* 1991, **45**(2):350-357.
5. Kirchhoff C: **Molecular characterization of epididymal proteins.** *Reviews of reproduction* 1998, **3**(2):86-95.
6. Bingle L SV, Bingle CD: **The putative ovarian tumour marker gene HE4 (WFDC2), is expressed in normal tissues and undergoes complex alternative splicing to yield multiple protein isoforms.** *Oncogene* 2002:2768-2773.
7. Bingle L, Cross SS, High AS, Wallace WA, Rassl D, Yuan G, Hellstrom I, Campos MA, Bingle CD: **WFDC2 (HE4): A potential role in the innate immunity of the oral cavity and respiratory tract and the development of adenocarcinomas of the lung.** *Respiratory research* 2006, **7**(1):61.
8. Nakajima K, Ono M, Radovic U, Dizdarevic S, Tomizawa SI, Kuroha K, Nagamatsu G, Hoshi I, Matsunaga R, Shirakawa T *et al.*: **Lack of whey acidic protein (WAP) four-disulfide core domain protease inhibitor 2 (WFDC2) causes neonatal death from respiratory failure in mice.** *Disease models & mechanisms* 2019, **12**(11).
9. Galgano MT, Hampton GM, Frierson HF, Jr.: **Comprehensive analysis of HE4 expression in normal and malignant human tissues.** *Mod Pathol* 2006, **19**(6):847-853.
10. LeBleu VS, Teng Y, O'Connell JT, Charytan D, Muller GA, Muller CA, Sugimoto H, Kalluri R: **Identification of human epididymis protein-4 as a fibroblast-derived mediator of fibrosis.** *Nat Med* 2013, **19**(2):227-231.
11. Raghu G, Richeldi L, Jagerschmidt A, Martin V, Subramaniam A, Ozoux ML, Esperet CA, Soubrane C: **Idiopathic Pulmonary Fibrosis: Prospective, Case-Controlled Study of Natural History and Circulating Biomarkers.** *Chest* 2018, **154**(6):1359-1370.
12. Raghu G, Collard HR, Egan JJ, Martinez FJ, Behr J, Brown KK, Colby TV, Cordier JF, Flaherty KR, Lasky JA *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.
13. Travis WD, Costabel U, Hansell DM, King TE, Jr., Lynch DA, Nicholson AG, Ryerson CJ, Ryu JH, Selman M, Wells AU *et al.*: **An official American Thoracic Society/European Respiratory Society**

- statement: Update of the international multidisciplinary classification of the idiopathic interstitial pneumonias.** *Am J Respir Crit Care Med* 2013, **188**(6):733-748.
14. Bohan A, Peter JB: **Polymyositis and dermatomyositis (first of two parts).** *N Engl J Med* 1975, **292**(7):344-347.
 15. **Preliminary criteria for the classification of systemic sclerosis (scleroderma).** Subcommittee for scleroderma criteria of the American Rheumatism Association Diagnostic and Therapeutic Criteria Committee. *Arthritis Rheum* 1980, **23**(5):581-590.
 16. Aletaha D, Neogi T, Silman AJ, Funovits J, Felson DT, Bingham CO, 3rd, Birnbaum NS, Burmester GR, Bykerk VP, Cohen MD *et al*: **2010 rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative.** *Annals of the rheumatic diseases* 2010, **69**(9):1580-1588.
 17. Ando M, Arima K, Yoneda R, Tamura M: **Japanese summer-type hypersensitivity pneumonitis. Geographic distribution, home environment, and clinical characteristics of 621 cases.** *Am Rev Respir Dis* 1991, **144**(4):765-769.
 18. Yoshizawa Y, Ohtani Y, Hayakawa H, Sato A, Suga M, Ando M: **Chronic hypersensitivity pneumonitis in Japan: a nationwide epidemiologic survey.** *J Allergy Clin Immunol* 1999, **103**(2 Pt 1):315-320.
 19. Collard HR, Ryerson CJ, Corte TJ, Jenkins G, Kondoh Y, Lederer DJ, Lee JS, Maher TM, Wells AU, Antoniou KM *et al*: **Acute Exacerbation of Idiopathic Pulmonary Fibrosis. An International Working Group Report.** *Am J Respir Crit Care Med* 2016, **194**(3):265-275.
 20. Edey AJ, Devaraj AA, Barker RP, Nicholson AG, Wells AU, Hansell DM: **Fibrotic idiopathic interstitial pneumonias: HRCT findings that predict mortality.** *Eur Radiol* 2011, **21**(8):1586-1593.
 21. Kanda Y: **Investigation of the freely available easy-to-use software 'EZ' for medical statistics.** *Bone marrow transplantation* 2013, **48**(3):452-458.
 22. Wijsenbeek M, Kreuter M, Olson A, Fischer A, Bendstrup E, Wells CD, Denton CP, Mounir B, Zouad-Lejour L, Quaresma M *et al*: **Progressive fibrosing interstitial lung diseases: current practice in diagnosis and management.** *Current medical research and opinion* 2019, **35**(11):2015-2024.
 23. Cottin V, Wollin L, Fischer A, Quaresma M, Stowasser S, Harari S: **Fibrosing interstitial lung diseases: knowns and unknowns.** *European respiratory review : an official journal of the European Respiratory Society* 2019, **28**(151).
 24. Jo HE, Glaspole I, Grainge C, Goh N, Hopkins PM, Moodley Y, Reynolds PN, Chapman S, Walters EH, Zappala C *et al*: **Baseline characteristics of idiopathic pulmonary fibrosis: analysis from the Australian Idiopathic Pulmonary Fibrosis Registry.** *Eur Respir J* 2017, **49**(2).
 25. Yokoyama A, Kohno N, Hamada H, Sakatani M, Ueda E, Kondo K, Hirasawa Y, Hiwada K: **Circulating KL-6 predicts the outcome of rapidly progressive idiopathic pulmonary fibrosis.** *Am J Respir Crit Care Med* 1998, **158**(5 Pt 1):1680-1684.
 26. Yokoyama A, Kondo K, Nakajima M, Matsushima T, Takahashi T, Nishimura M, Bando M, Sugiyama Y, Totani Y, Ishizaki T *et al*: **Prognostic value of circulating KL-6 in idiopathic pulmonary fibrosis.** *Respirology (Carlton, Vic)* 2006, **11**(2):164-168.

27. Okamoto T, Fujii M, Furusawa H, Tsuchiya K, Miyazaki Y, Inase N: **The usefulness of KL-6 and SP-D for the diagnosis and management of chronic hypersensitivity pneumonitis.** *Respir Med* 2015, **109**(12):1576-1581.
28. Choi SI, Jang MA, Jeon BR, Shin HB, Lee YK, Lee YW: **Clinical Usefulness of Human Epididymis Protein 4 in Lung Cancer.** *Ann Lab Med* 2017, **37**(6):526-530.
29. Hellström I RJ, Hayden-Ledbetter M, Ledbetter JA, Schummer M, McIntosh M, Drescher C, Urban N, Hellström KE.: **The HE4 (WFDC2) protein is a biomarker for ovarian carcinoma.** *Cancer Res* 2003, **63**(13):3695-3700.
30. Nagy B, Jr., Krasznai ZT, Balla H, Csoban M, Antal-Szalmas P, Hernadi Z, Kappelmayer J: **Elevated human epididymis protein 4 concentrations in chronic kidney disease.** *Annals of clinical biochemistry* 2012, **49**(Pt 4):377-380.
31. Bolstad N, Oijordsbakken M, Nustad K, Bjerner J: **Human epididymis protein 4 reference limits and natural variation in a Nordic reference population.** *Tumour Biol* 2012, **33**(1):141-148.
32. Bingle CD, Vyakarnam A: **Novel innate immune functions of the whey acidic protein family.** *Trends Immunol* 2008, **29**(9):444-453.
33. Hewlett JC, Kropski JA, Blackwell TS: **Idiopathic pulmonary fibrosis: Epithelial-mesenchymal interactions and emerging therapeutic targets.** *Matrix biology : journal of the International Society for Matrix Biology* 2018, **71-72**:112-127.
34. Raghu G, Remy-Jardin M, Myers JL, Richeldi L, Ryerson CJ, Lederer DJ, Behr J, Cottin V, Danoff SK, Morell F *et al.*: **Diagnosis of Idiopathic Pulmonary Fibrosis. An Official ATS/ERS/JRS/ALAT Clinical Practice Guideline.** *American Journal of Respiratory and Critical Care Medicine* 2018, **198**(5):e44-e68.
35. Strieter RM: **Pathogenesis and natural history of usual interstitial pneumonia: the whole story or the last chapter of a long novel.** *Chest* 2005, **128**(5 Suppl 1):526s-532s.

Tables

Table 1 Patient characteristics

	Derivation cohort			Validation cohort		
	PF-ILD	HV	<i>P</i> value	PF-ILD	Control	<i>P</i> value
No. of cases, n	34	40		39	40	
Age, years	68.5 (62.2-73.7)	50.5 (47-55.3)	< 0.01	66 (59-82)	70 (62.5-85)	0.44
Gender, n			0.56			0.21
Male	19	25		25	20	
Female	15	15		14	20	
Smoking history						0.01
Ex-smoker	20	ND		16	25	
Never smoker	14	ND		23	11	
Pack-years	26.4 ± 29.3	ND		16.3 ± 27.5	26.5 ± 36.5	0.17
Laboratory data						
KL-6 IU/ml	2140 ± 2163	ND		1360 ± 641	ND	
SP-D ng/ml	348 ± 314	ND		367 ± 275	ND	
eGFR ml/min/1.73m ²	85.6 ± 24.4	ND		73.7 ± 13.7	78.1 ± 16.8	0.22
Pulmonary function test, % predicted						
FVC	66.6 ± 17.7 [§]	ND		65.1 ± 17.0	ND	
FEV1	72.1 ± 16.4 [§]	ND		70.5 ± 17.7	ND	
DLCO/VA	56.6 ± 14.3 ^{§§}	ND		73.7 ± 21.3 ^{§§§}	ND	
Etiology of PF-ILDs						
CHP	16	0		7	0	
IIPs	11	0		28	0	
IPF	6			13		
non-IPF	5			15		
CTD	6	0		4	0	
SSc	2			1		
RA	2			1		
DM	2			2		
Asbestosis	1	0		0	0	

Values are given as number or the median (25th and 75th percentile). A comparison between PF-ILDs groups was performed using Welch's t test, the chi-square(χ^2) test, and Mann-Whitney's U test as appropriate.

Abbreviations; KL-6: Krebs von den Lungen 6; SP-D: surfactant protein D; eGFR: estimated glomerular filtration rate; FVC: forced vital capacity; FEV1: forced expiratory volume; DLCO: diffusing capacity for carbon monoxide; VA: alveolar volume; PF-ILD: progressive-fibrosing interstitial lung disease; CHP: chronic hypersensitivity pneumonitis; IIPs: idiopathic interstitial pneumonias; IPF: idiopathic pulmonary fibrosis; CTD: collagen tissue disease; RA: Rheumatoid Arthritis; SSc: Systemic sclerosis; DM: dermatomyositis; ND: no data

§: n = 33, §§: n = 27, §§§: n = 21

Table 2 Correlation between serum HE4 levels and clinical variables

	Derivation cohort	
	r	P value
Age years	0.49	< 0.003
Pack-years	0.14	0.44
Laboratory date		
eGFR ml/min/1.73m ²	-0.27	0.12
KL-6 IU/ml	< -0.01	1.00
SP-D ng/ml	0.30	0.09
Pulmonary function test % predicted		
FVC	-0.18	0.32
FEV1	-0.18	0.31
DLCO/VA	-0.28	0.16
Score of HRCT Findings		
Reticulation	0.11	0.55
Honeycombing	0.41	0.015
Centrilobular nodules	-0.29	0.10
Ground-glass opacity	0.16	0.38
Consolidation	0.14	0.45
Emphysema	-0.21	0.24
TBE	0.29	0.097

Bold value indicates statistically significance (P value < 0.05).

eGFR: estimated glomerular filtration rate; KL-6: Krebs von den Lungen 6; SP-D: surfactant protein D; FVC: forced vital capacity; FEV1: forced expiratory volume; DLCO: diffusing capacity for carbon monoxide; VA: alveolar volume; HRCT: high-resolution computed tomography; TBE: traction bronchiectasis.

Table 3. Comparison of candidate risk factors between surviving and non-surviving patients with PF-ILD

	Derivation cohort			
	n	non-survivors	survivors	P value
		9	25	
Age, n (%)				0.11
< 65		1 (11.1)	11 (44)	
≥ 65		8 (88.9)	14 (56)	
Gender, n (%)				0.70
Male		6 (66.7)	13 (52)	
Female		3 (33.3)	12 (48)	
Smoking history, n (%)				0.25
Yes		7 (77.8)	13 (52)	
No		2 (22.2)	12 (48)	
Laboratory test, mean ± SD				
HE4 pmol/l		221±70.5	156±62.1	0.013
KL-6 U/mL		2320±2773	2076±1964	0.78
SP-D ng/mL		382±356	337±307	0.73
eGFR ml/min/1.73m ²		79.9±27.5	87.7±23.5	0.42
Respiratory function test				
%predicted FVC<80% (n = 26)		7 (77.8)	19 (76)	0.65
%predicted FVC≥80% (n =7)		1 (11.1)	6 (24)	
Decline in FVC from baseline (ml/year) (n = 22)		-337±151 (n = 4)	-100±532 (n = 18)	0.12
Etiology of PF-ILD, n (%)				0.64
IPF		2 (22.2)	4 (16)	
Non-IPF		7 (77.8)	21 (84)	
Radiological pattern of ILD, n (%)				0.55
UIP		0	4 (16)	
non-UIP		9 (100)	21 (84)	

Values are given as number (%) or the mean ± SD.

Bold value indicates statistically significance (*P* value < 0.05).

Abbreviations; eGFR: estimated glomerular filtration rate; HE4: human epididymis protein 4; KL-6: Krebs von den Lungen 6; SP-D: surfactant protein D; FVC: forced vital capacity; IPF: idiopathic pulmonary fibrosis; UIP: usual interstitial pneumonia.

Table 4 Cox proportional hazards model for mortality

Variables	Derivation cohort			Validation cohort		
	Hazard ratio	95% CI	<i>P</i> value	Hazard ratio	95% CI	<i>P</i> value
Univariate						
Age year (≥ 65)	6.02	0.75-48.4	0.09			
Gender (male)	1.74	0.43-6.96	0.43			
Smoking history (smoker)	3.30	0.68-15.9	0.14			
HE4 (>238pmol/l)	6.54	1.71-25	0.006			
KL-6 IU/ml (≥ 1000)	0.67	0.22-2.01	0.47			
SP-D ng/ml (≥ 200)	2.02	0.41-10.0	0.39			
FVC % predicted (<80%)	3.22	0.39-26.4	0.28			
Etiology of PF-ILD (IPF)	1.03	0.21-4.97	0.97			
UIP pattern on HRCT	N/A*	N/A*	0.99			
Multivariate						
Age year (≥ 65)	12.0	1.31-109	0.027	4.51	0.76-26.9	0.098
HE4 (>238pmol/l)	7.27	1.56-34.0	0.012	44.3	4.19-468	0.002
Smoking history (smoker)	5.26	0.95-29.1	0.057	1.76	0.42-7.43	0.44

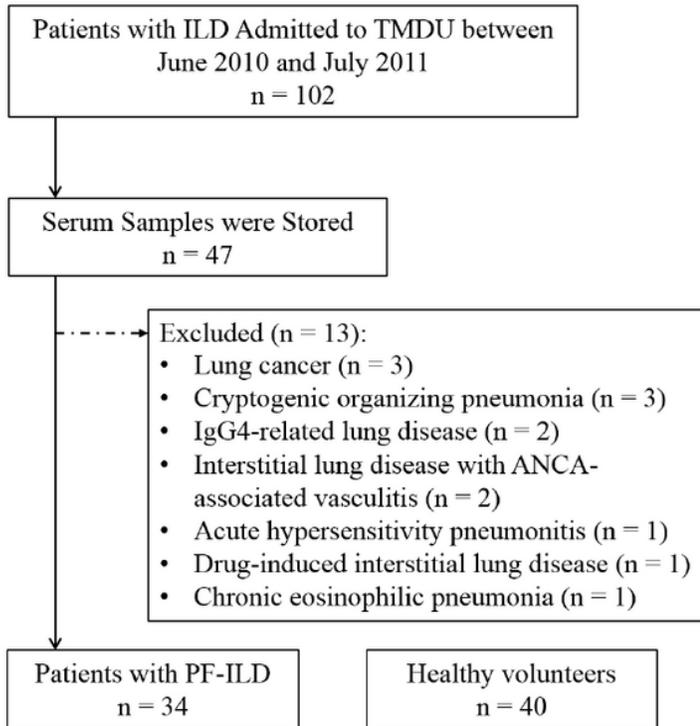
Bold values indicate statistically significance (*P* value < 0.05).

HE4: human epididymis protein 4; KL-6: Krebs von den Lungen 6; SP-D: surfactant protein D; FVC: forced vital capacity; eGFR: estimated glomerular filtration rate.

* Unanalyzable data because of none of non-survivors with UIP pattern

Figures

A) Derivation cohort



B) Validation cohort

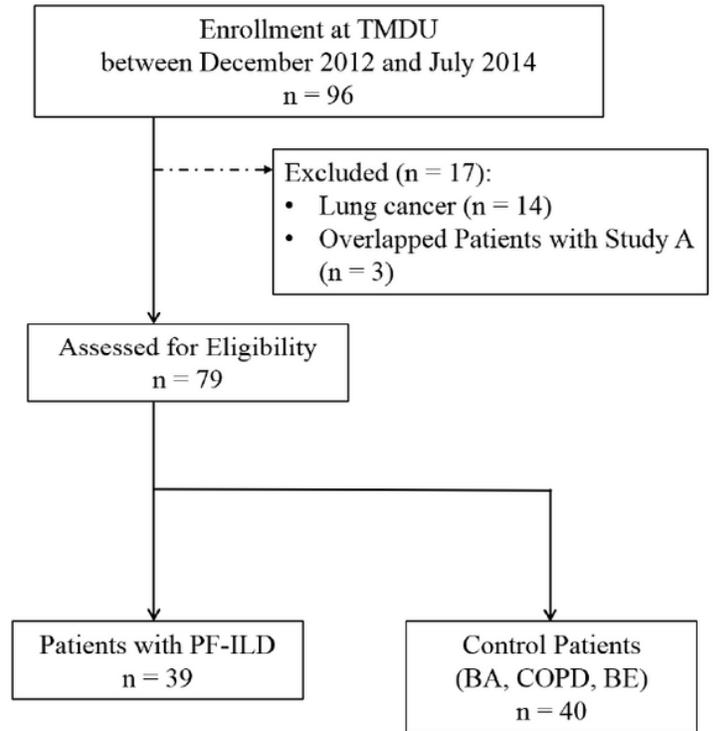


Figure 1

CONSORT flow diagram for study population.

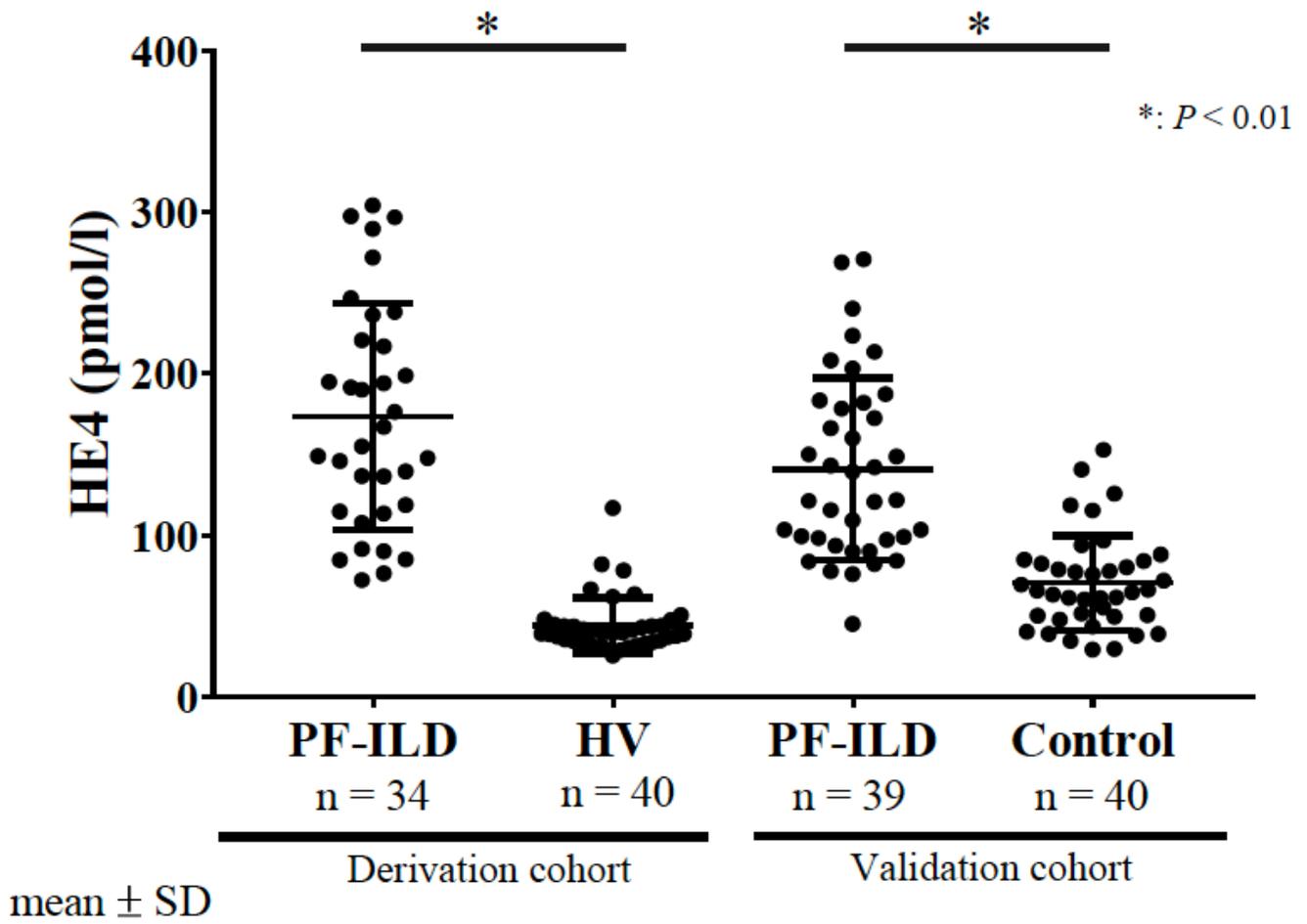
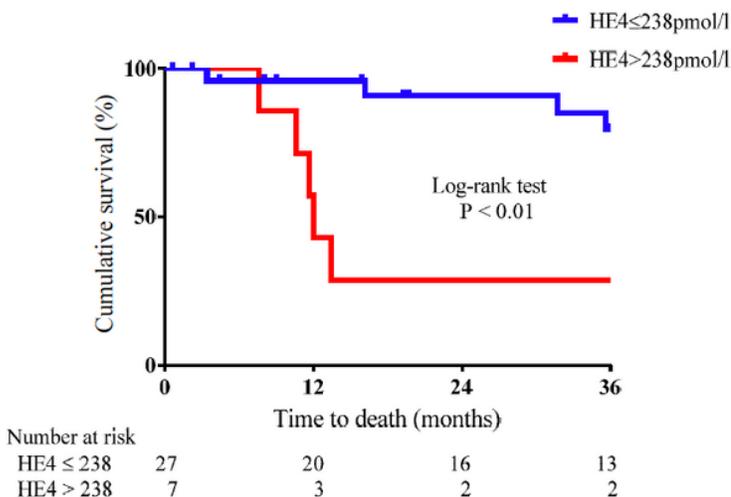


Figure 2

Serum human epididymis protein 4 (HE4) levels in patients with progressive fibrosing interstitial lung disease and control subjects. Solid lines and bars indicate group means \pm SD. *: $P < 0.01$.

A. Derivation cohort



B. Validation cohort

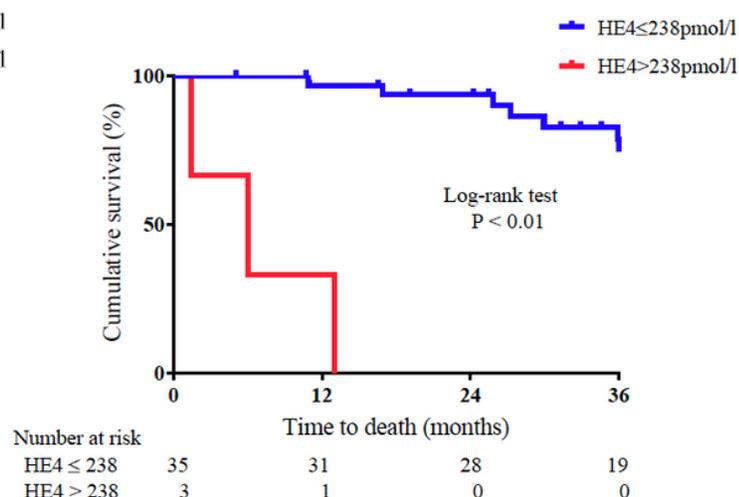


Figure 3

Kaplan-Meier curves for survival between the higher and lower HE4 groups in the derivation cohort (A) and the validation cohort (B).



Figure 4

Expression of human epididymis protein 4 (HE4) in lung tissues. A, D) Normal control lung tissue shows no HE4 expression in bronchiolar epithelium and lung parenchyma. B, E) Lung tissue of usual interstitial pneumonia (UIP) shows HE4 expression that is detected in metaplastic bronchiolized epithelium lining regions of honeycombing and bronchiolar epithelium. C, F) Lung tissue of nonspecific interstitial pneumonia (NSIP) shows HE4 expression in bronchiolar epithelium.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [HE4SupplementaryTable.docx](#)
- [SupplementaryFigureS1.pdf](#)