

Human epididymis protein 4 as a new diagnostic biomarker for rheumatoid arthritis-associated interstitial lung disease

T. Lin¹, S. Xu¹, Y. Wang², X. Nian³, X. Shan¹, T. Jiang¹, M. Qiu¹

¹Department of Respiratory and Critical Care Medicine, ²Department of Rheumatology and Immunology, ³Department of Gastroenterology, Yantai Yuhuangding Hospital, affiliated with the Medical College of Qingdao, Yantai, Shandong, China.

Abstract

Objective

This study aimed to evaluate the role of human epididymis protein 4 (HE4) in the diagnosis and determination of the severity of interstitial lung disease (ILD) in rheumatoid arthritis (RA) patients.

Methods

HE4 levels in peripheral blood (PB) and bronchoalveolar lavage fluid (BALF) samples were determined via electrochemiluminescence immunoassays in 102 RA patients (46 patients with ILD and 56 patients without ILD) and 51 healthy controls (HCs).

Results

Serum HE4 levels were significantly higher in RA-ILD patients (141.8 ± 65.92 pmol/l) than those in the RA-no ILD patients (82.67 ± 26.17 pmol/l) and healthy controls (35.72 ± 7.6 pmol/l) ($p < 0.0001$). Consistent with serum HE4 levels, BALF HE4 levels were significantly higher in RA-ILD patients (637.6 ± 154.9 pmol/l) than those in the RA-no ILD patients (427.3 ± 111.2 pmol/l) and healthy controls (206.9 ± 30.46 pmol/l) ($p < 0.0001$). In RA-ILD patients, HE4 levels were positively correlated with HRCT (high-resolution computed tomography) fibrosis scores, whereas a significant inverse relationship was found between HE4 levels and lung function parameters (such as, diffusion capacity of the lung for carbon monoxide (DLCO)). The logistic regression analysis showed that high levels of BALF HE4 (≥ 595 pmol/l) were associated with RA-ILD (odds ratio [OR] = 8.09; 95% confidence interval [CI] = 1.317–49.682; $p = 0.024$).

Conclusion

Serum and BALF HE4 levels were elevated in RA-ILD patients and strongly associated with the severity of ILD, thus supporting their potential clinical value as a new diagnostic aid for patients with RA-ILD.

Key words

human epididymis protein 4, interstitial lung disease, rheumatoid arthritis

Tiantian Lin, BSc*
 Shufen Xu, BSc*
 Yutie Wang, MD*
 Xueyuan Nian, MD
 Xiaoxi Shan, MD
 Tingshu Jiang, PhD**
 Meihua Qiu, MD**

*These authors contributed equally to this work and are co-first authors.

**These authors contributed equally to this work and are co-corresponding authors.

Please address correspondence to:
 Meihua Qiu,
 Department of Respiratory
 and Critical Care Medicine,
 Yantai Yuhuangding Hospital,
 Affiliated with the Medical
 College of Qingdao,
 20 Yuhuangding East Road,
 Yantai, Shandong 264200, China.
 E-mail: qiumeihua901013@163.com

and to:

Tingshu Jiang
 E-mail: huxi901013@163.com

Received on December 7, 2021; accepted
 in revised form on February 28, 2022.

© Copyright CLINICAL AND
 EXPERIMENTAL RHEUMATOLOGY 2022.

Funding: this study was supported by grants from the Science and Technology Program of Yantai City (no. 2020YD024).
 Competing interests: none declared.

Introduction

Rheumatoid arthritis (RA) is a chronic, systemic, inflammatory disease that involves the synovial joints and it can lead to progressive bone destruction and eventual joint deformity if left untreated (1). RA patients may have extra-articular involvement associated with impairment in physical function, higher morbidity, and premature mortality (2). RA may affect the lung interstitium, airways, and pleurae, whereas pulmonary vascular involvement is less frequent (3). Among those involvements, clinical and preclinical interstitial lung disease (ILD) is the most common and can be detected in 60% of patients with RA (4-7). RA-associated interstitial lung disease (RA-ILD) has a heterogeneous clinical presentation and disease course. ILD is still a major cause of mortality, with a median survival of only 3-9 years after diagnosis (4, 7, 8).

Better methods of early diagnosis and monitoring of disease progression in RA-ILD patients are needed to improve the prognosis of patients. At present, the diagnosis of RA-ILD mainly depends on the use of HRCT, but costs and ionising radiation may limit its use in clinical practice. The discovery of new non-invasive biomarkers that could help improve the prospective evaluation of patients with RA-ILD is urgently needed.

Human epididymis protein 4 (HE4) is encoded by the whey acidic protein four-disulfide domain 2 (WFDC2) gene, which is known as a tumour marker in epithelial ovarian cancers (9), lung cancer (10, 11), and endometrial carcinomas (12, 13). Additionally, it has been found to be a potential biomarker for the detection of renal fibrosis and cystic fibrosis, as well as being a potential target to inhibit fibrosis (14-16). HE4 is elevated in patients with chronic kidney disease, and its concentration is associated with severities in kidney function and renal fibrosis (14). HE4 suppresses the activity of multiple proteases, including serine proteases and matrix metalloproteinases, and it specifically inhibits their capacity to degrade type I collagen, which suggests a functional role of HE4 in the pathogenesis of fibrosis. In lung disease, the elevated

expression of HE4 mRNA was also described in cystic fibrosis (CF) lung biopsy specimens compared with non-CF control samples, and HE4 serum levels were positively correlated with the overall severity of CF (14). In addition, serum HE4 levels were reportedly higher in patients with IPF than in patients with no respiratory disease (17). In patients with progressive fibrosing interstitial lung disease (PF-ILD), serum HE4 levels were shown to have a positive correlation with the extent of honeycombing on chest high-resolution computed tomography (HRCT) (18). Moreover, higher HE4 levels were associated with an elevated mortality risk, which showed that serum HE4 levels may serve as a potential predictive biomarker for prognosis in PF-ILD. Moreover, a recent study showed that HE4 levels in serum and bronchoalveolar lavage fluid (BALF) were markedly increased in patients with systemic sclerosis-related interstitial lung disease and had a positive correlation with ILD severity (19).

However, little is known about the role of HE4 as a biomarker for the diagnosis of RA-ILD. The purpose of this study was to evaluate the predictive value of HE4 and to examine the correlation between HE4 and ILD severity.

Methods

Study population

We recruited 56 consecutive Chinese Han patients with newly diagnosed RA-no-ILD, 46 patients with RA-ILD, and 51 healthy controls (HCs) at the Yantai Yuhuangding Hospital, Qingdao Medical University, China (Table I). The diagnosis of RA was based on the 2010 American College of Rheumatology/European League Against Rheumatism (ACR/ELAR) criteria (20). Patients were excluded from this study if they presented with signs of other pulmonary diseases (such as pulmonary infection, pulmonary tuberculosis, chronic obstructive pulmonary disease, and ILD caused by pneumoconiosis and other causes), a history of malignancy, a history of chronic kidney disease, or a recent clinically significant infection (such as human immunodeficiency virus or viral hepatitis). The diagnosis of ILD was

based on clinical findings, pulmonary function tests (PFTs), and HRCT scans at entry. The common clinical symptoms of RA-ILD include chronic dry cough, exertional dyspnoea, fatigue, and generalized weakness. Healthy controls included volunteers who were recruited and who were age matched with the case groups without evidence of any diseases. The study protocol (no. 2018-26) was approved by the Ethics Committee of Yantai Yuhuangding Hospital, and all the participants provided written informed consent.

Pulmonary function tests

Screening spirometry was performed with the Master Screen PFT System (JAEGGER, Germany), in accordance with American Thoracic Society/European Respiratory Society recommendations (21). The forced expiratory volume in 1 second (FEV1), forced vital capacity (FVC), FEV1/FVC ratio, and diffusion capacity of the lung for carbon monoxide (DLCO) were recorded. These values were expressed as a percentage of the predicted value compared with individuals of similar sex, age, weight, and height. A restrictive pattern was defined as a FVC of <80% of that predicted in the absence of concomitant obstructive abnormalities.

Chest high-resolution computed tomography analysis

All the images were independently evaluated by two experienced thoracic radiologists who were blinded to the clinical data. Discordant interpretations were resolved via a consensus between both radiologists. HRCT results were analysed for ILD scoring. Scoring was performed in each of 3 lung zones: the upper zone (lung apex to aortic arch), the middle zone (aortic arch to inferior pulmonary veins), and the lower zone (inferior pulmonary veins to diaphragm). HRCT scans were semi-quantitatively graded based on lung involvement by using a Likert scale (0=absent, 1=1–25%, 2=26–50%, 3=51–75%, and 4=76–100%) for the extent of 4 categories of parenchymal abnormality (ground glass opacity: hazy parenchymal opacity in the absence of reticular opacity or architectural distortion; lung

Table I. Demographics and clinical data of the study participants.

	RA-no-ILD (n=56)	RA-ILD (n=46)	HCs (n=51)	<i>p</i> -value
Age (years)	58.29 ± 8.83	63.35 ± 8.16	57.16 ± 4.11	0.000
Female, n (%)	29/27	31/15	26/25	0.179
Smoking history, n (%)	23 (40.1)	12 (26.1)	27 (52.9)	0.027
Current smokers, n (%)	8 (14.3)	10(21.7)	12 (23.5)	0.441
BMI (kg/m ²)	23.86 ± 2.9	24.36 ± 2.73	24.68 ± 2.95	0.573
RA disease duration (years)	2 (0.08-40)	3 (0.5-10)	-	0.661
DAS28 (units)	5.38 ± 1.18	5.67± 1.35	-	0.213
RF titre (IU/ml)	93.3 (9.2-1420)	137 (9.2-7560)	-	0.000
RF positive, n (%)	46 (82.1)	43 (93.5)	-	0.088
Anti-CCP antibody titre (IU/ml)	96.8 (1-200)	200(1-200)	-	0.028
Anti-CCP antibody positive, n (%)	54 (96.4)	45 (97.8)	-	0.678
Anti-CCP antibody high titre positive, n (%)	23 (41.1)	28(60.87)	-	0.047
ANA positive, n (%)	40 (71.43)	39(84.78)	-	0.108
ESR (mm/h)	41.96 ± 25.54	49.13 ± 24.76	-	0.867
CRP (mg/dl)	32.31 ± 26.82	55.64 ± 108.96	-	0.025

RA-ILD: rheumatoid arthritis-associated interstitial lung disease; HCs: healthy controls; BMI: body mass index; DAS: disease activity score; RF: rheumatoid factor; anti-CCP: anti-cyclic citrullinated peptide; ANA: antinuclear antibody; ESR: erythrocyte sedimentation rate; CRP: C-reactive protein.

fibrosis: reticular opacification, traction bronchiectasis, and bronchiolectasis; honeycombing: clustered air-filled cysts with dense walls; and emphysema: lucencies or cysts without walls) (22). This scoring system is based on that reported by Kazerooni *et al.* (23). The scores for each lung zone were averaged for data analysis.

Sample collection and measurements of HE4

Serum samples were obtained from each subject. BALF specimens were obtained from 38 RA-ILD patients and 30 RA-no-ILD patients according to a standardised protocol (24). Twenty-one subjects underwent routine healthy examinations, received diagnostic bronchoscopies and showed normal BAL cytology measurements. The BAL cell differentials of each group are shown in Table II. The obtained serum and BALF samples were centrifuged immediately at room temperature. All the serum samples and BALF supernatant samples were stored at -80°C prior to use. The concentration of HE4 was determined via an electrochemiluminescence immunoassay (ECLI, Roche Diagnostics GmbH, Mannheim, Germany) in the laboratory at Yantai Yuhuangding Hospital according to the manufacturer's instructions.

Statistical analysis

All the analyses were performed with SPSS for Windows v. 17.0 (Chicago, Illinois, USA). Continuous variables are summarised as medians (interquartile ranges [IQRs]) or means ± standard deviations (SDs) and categorical data with frequency (percentages). Groups were compared by using an analysis of variance, Student's *t*-test, Wilcoxon rank-sum test, or a χ^2 test (as appropriate). Serum and BALF HE4 levels of the different groups were compared by using an analysis of variance. Correlations were assessed by using a Pearson correlation test or a Spearman's rank test. The receiver-operating characteristic (ROC) curve was used to analyse the value of HE4 in the diagnosis of RA-ILD. Sensitivity, specificity, and positive and negative predictive values were calculated based on cut-off levels that were identified in the ROC analyses that maximised sensitivity and specificity. A multivariate logistic regression model was built to identify risk factors for ILD in RA patients. A *p*-value <0.05 was considered to be statistically significant.

Results

Demographic characteristics

The characteristics of 56 patients with RA-no-ILD, 46 patients with RA-ILD,

and 51 healthy controls are summarised in Table I. Twenty-seven RA-ILD patients (58.7%) showed respiratory symptoms, of which shortness of breath was most frequent (18/27, 66.7%). The mean age of the RA-ILD patients was 63.35 ± 8.16 years and that of RA-no-ILD cases was 58.29 ± 8.83 years. RA-ILD patients were older than RA no-ILD patients and healthy controls ($p < 0.05$). RA-ILD patients had higher rheumatoid factor (RF) and anti-cyclic citrullinated peptide (ACCP) antibody titres and higher C-reactive protein (CRP) levels than RA no-ILD patients ($p < 0.05$), whereas other general biological indices, such as body mass index (BMI), the duration of the disease process, disease activity score (DAS)28, antinuclear antibody (ANA) positivity, and erythrocyte sedimentation rate (ESR), showed no significant differences between the two groups. Laboratory findings are shown in Table II.

Increased serum and BALF HE4 levels in RA-ILD patients

We first investigated the HE4 levels in serum and BALF among the different groups. As shown in Fig. 1a, the serum HE4 levels were significantly elevated in RA-ILD patients (141.8 ± 65.9 pmol/l) compared with RA-no-ILD patients (82.67 ± 26.2 pmol/l) and healthy control subjects (35.72 ± 7.6 pmol/l) ($p < 0.0001$). In addition, the serum HE4 levels were also higher in RA-no-ILD patients (82.67 ± 26.2 pmol/l) than in healthy controls (35.72 ± 7.6 pmol/l) ($p < 0.0001$). As shown in Fig. 1b, BALF HE4 levels were also significantly elevated in RA-ILD patients (637.6 ± 154.9 pmol/l) compared with RA-no-ILD patients (427.3 ± 111.2 pmol/l) and healthy control subjects (206.9 ± 30.46 pmol/l) ($p < 0.0001$). In addition, the BALF HE4 levels were also higher in RA-no-ILD patients (427.3 ± 111.2 pmol/l) than in healthy controls (206.9 ± 30.46 pmol/l) ($p < 0.0001$). Moreover, a significant relationship was observed between the serum and BALF HE4 levels ($r = 0.7606$, $p < 0.0001$, Fig. 2).

Increased HE4 levels correlated with ILD severity in RA-ILD patients

We also explored the correlation be-

Table II. Laboratory findings in patients and healthy controls.

	RA-ILD (n=38)	RA-no-ILD (n=30)	HCS (n=21)	p-value
PaO ₂ (mmHg)	73.40 ± 10.31	95.14 ± 3.15	97.65 ± 1.51	0.000
Pulmonary function test				
FEV1%, predicted	92.42 ± 20.56	94.46 ± 3.49	95.36 ± 3.47	0.490
FVC%, predicted	86.31 ± 18.96	95.20 ± 3.39	92.74 ± 3.21	0.000
DLCO%, predicted	72.88 ± 22.23	88.80 ± 4.12	91.17 ± 3.21	0.000
ILD score				
GGO score	0.25 (0-3.33)	-	-	
Fibrosis score	1 (0.16-3.83)	-	-	
Honeycombing score	0 (0-3.17)	-	-	
Emphysema score	0 (0-2.33)	-	-	
BAL cellular profile				
Macrophage, %	86.32 ± 4.66	89.8 ± 3.77	90.28 ± 3.12	0.001
Lymphocyte, %	7.13 ± 3.04	7.31 ± 3.53	6.12 ± 2.8	0.394
Neutrophil, %	6.11 ± 2.52	2.34 ± 1.80	3.00 ± 0.91	0.000
Eosinophil, %	0.85 ± 0.5	1.07 ± 0.92	0.60 ± 0.45	0.057
Serum HE4 (pmol/l)	141.8 ± 65.92	82.67 ± 26.17	35.72 ± 7.6	0.000
BALF HE4 (pmol/l)	637.6 ± 154.9	427.3 ± 111.2	206.9 ± 30.46	0.000

RA-ILD: rheumatoid arthritis-associated interstitial lung disease; HCS: healthy controls; PaO₂: partial pressure of arterial oxygen; FEV1: forced expiratory volume in the first second; FVC: forced vital capacity; DLCO: diffusing capacity of the lung for carbon monoxide; GGO: ground glass opacity; BALF: bronchoalveolar lavage fluid; HE4: human epididymis protein 4.

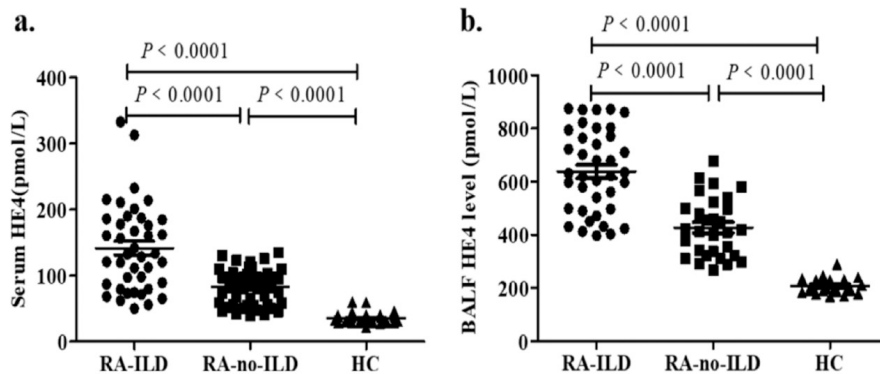


Fig. 1. Serum (a) and BALF (b) HE4 levels in patients with RA-ILD or RA-no-ILD and HCs. BALF: bronchoalveolar lavage fluid; HE4: human epididymis protein 4; RA-ILD: rheumatoid arthritis-associated interstitial lung disease; HCs healthy controls.

tween HE4 levels and ILD severity in RA patients by using a Spearman analysis. We found that HE4 levels in serum and BALF were apparently correlated with HRCT fibrosis scores in the RA-ILD group (Fig. 3a: $r = 0.5256$, $p = 0.0004$; Fig. 3b: $r = 0.4063$, $p = 0.0114$). In addition, we evaluated the relationship between HE4 levels and PFTs. In the RA-ILD group, the HE4 levels in serum and BALF showed a significant inverse correlation with DLCO% pred (Fig. 4a: $r = -0.3848$, $p = 0.0393$; Fig. 4b: $r = -0.4472$, $p = 0.022$). Furthermore, the relationship between HE4 levels and the values of FVC% pred was analysed in RA-ILD patients; although the p -value was not

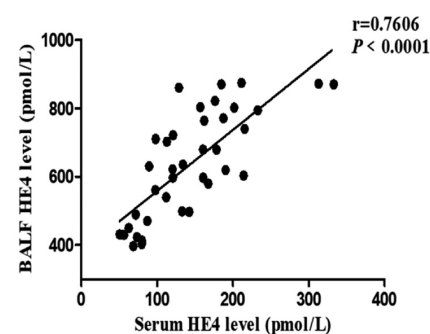


Fig. 2. Associations between HE4 in serum and BALF. BALF: bronchoalveolar lavage fluid; HE4: human epididymis protein 4.

significant, an inverse tendency was found (Fig. 4c: $r = -0.1655$, $p = 0.3909$; Fig. 4d: $r = -0.3015$, $p = 0.1344$).

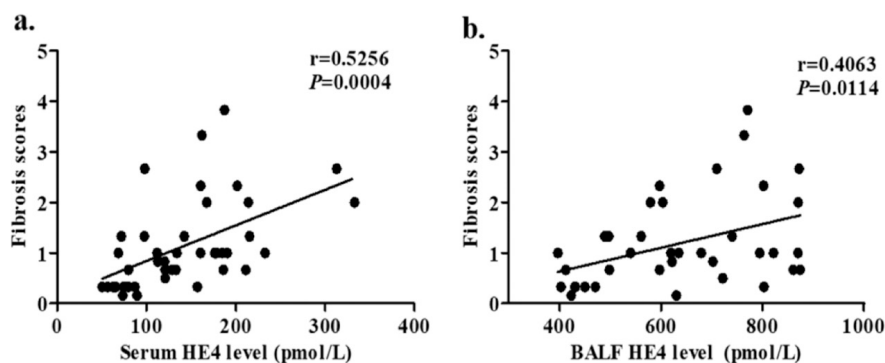


Fig. 3. Associations between HE4 and HRCT fibrosis scores in patients with RA-ILD. **a.** Serum HE4 in vs. fibrosis scores. **b.** BALF HE4 vs. fibrosis scores. HE4: human epididymis protein 4; HRCT: high-resolution computed tomography; BALF: bronchoalveolar lavage fluid.

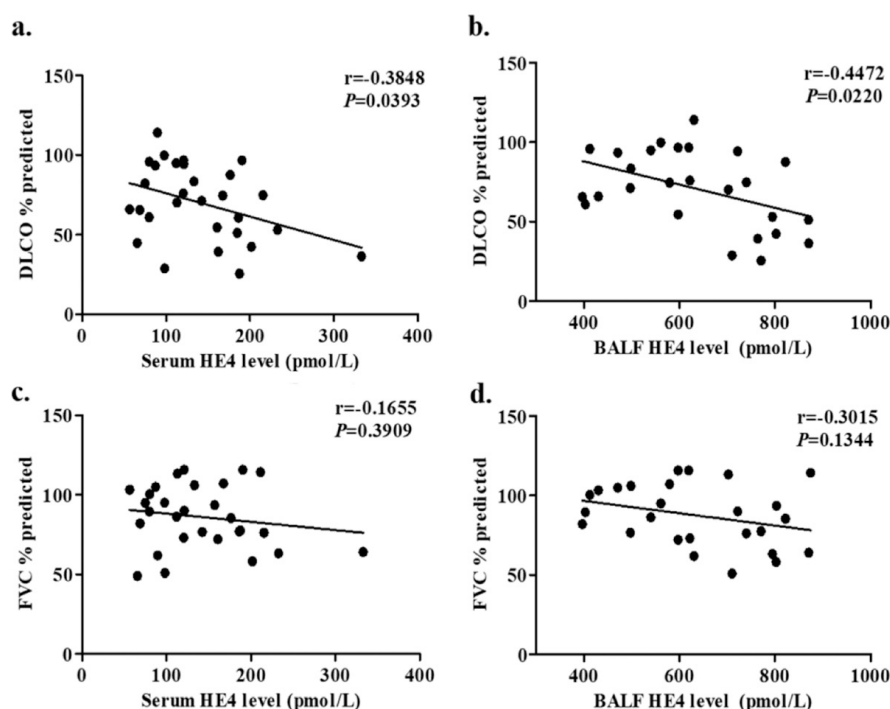


Fig. 4. Associations between HE4 and lung function values in patients with RA-ILD. **a.** Serum HE4 vs. DLCO % predicted. **b.** BALF HE4 vs. DLCO % predicted. **c.** Serum HE4 vs. FVC% predicted. **d.** BALF HE4 vs. FVC% predicted. HE4: human epididymis protein 4; DLCO: diffusing capacity of the lung for carbon monoxide; FVC: forced vital capacity; BALF bronchoalveolar lavage fluid.

Clinical value of HE4 in the diagnosis of ILD in RA patients

The predictive power of serum and BALF HE4 for detecting the presence of ILD in RA patients was calculated by Receiver operating curve (ROC)/area under the curve (AUC) through plotting sensitivity against specificity. The predictive ROC/AUC of serum HE4 predictive power was 0.782 (95% CI: 0.683–0.881), with a sensitivity and specificity at 65.9% and 86%, respectively (the cut-off value was 111.1 pmol/l; Fig. 5); The predictive ROC/AUC of BALF HE4 predic-

tive power was 0.856 (95% CI: 0.759–0.952), with a sensitivity and specificity at 63.1% and 90.0%, respectively (the cut-off value was 595 pmol/l; Fig. 5). We further identified the variables that were associated with RA-ILD via a logistic regression analysis (Table III). Age, sex, smoking history, BMI, disease duration, RF titre, anti-CCP antibody titre, ESR, and CRP were included in this model. After adjusting for the mutual influence of variables, the multivariate logistic regression analysis demonstrated that high levels of BALF HE4 (≥ 595 pmol/l) (odds ratio

[OR] = 8.09; 95% CI=1.317–49.682; $p=0.024$) were significantly associated with RA-ILD, which meant that high levels of BALF HE4 may increase the morbidity of ILD in RA (Table III). Although the p -value was not <0.05 , there was a trend that high levels of serum HE4 (≥ 111.1 pmol/l) (OR=5.082; 95% CI=0.986–26.188; $p=0.052$) were associated with the development of RA-ILD. If the included sample was increased, the p -value of the results may be significant.

Discussion

To our knowledge, this is the first study to research the role of HE4 in RA patients. In this study, the HE4 levels of serum and BALF in RA-ILD patients were significantly higher than those in RA-no-ILD patients. Furthermore, the ROC curve further demonstrated that HE4 may be a potential biomarker in discriminating RA-ILD patients from RA patients. Moreover, the HE4 level was correlated with impairment in two parameters for ILD severity, including the decrease in DLCO pred% and higher fibrosis scores for ILD involvement identified in the HRCT. These results emphasise the potential for HE4 to be used as a novel diagnostic aid for RA-ILD.

HE4 is also known as the WFDC2 protein, which is a member of the whey acidic proteins family that displays a proteinase inhibitor function, and HE4 specifically inhibits the degradation of type I collagen by suppressing Prss35 and Prss23 serin protease activity (25, 26). The protein was originally identified as a transcript that is exclusively expressed in the epididymis and was proposed as being a specific marker for this tissue (25). HE4 has widely served as a routine clinical biomarker in the diagnosis and follow-up of patients with epithelial ovarian cancer (27). The United States Food and Drug Administration has approved HE4 as a biomarker to monitor the recurrence and the progression of ovarian cancer (28). HE4 has been found to be expressed in a number of normal human tissues outside of the male reproductive system, including the trachea, lung and nasal epithelium and was also found in a

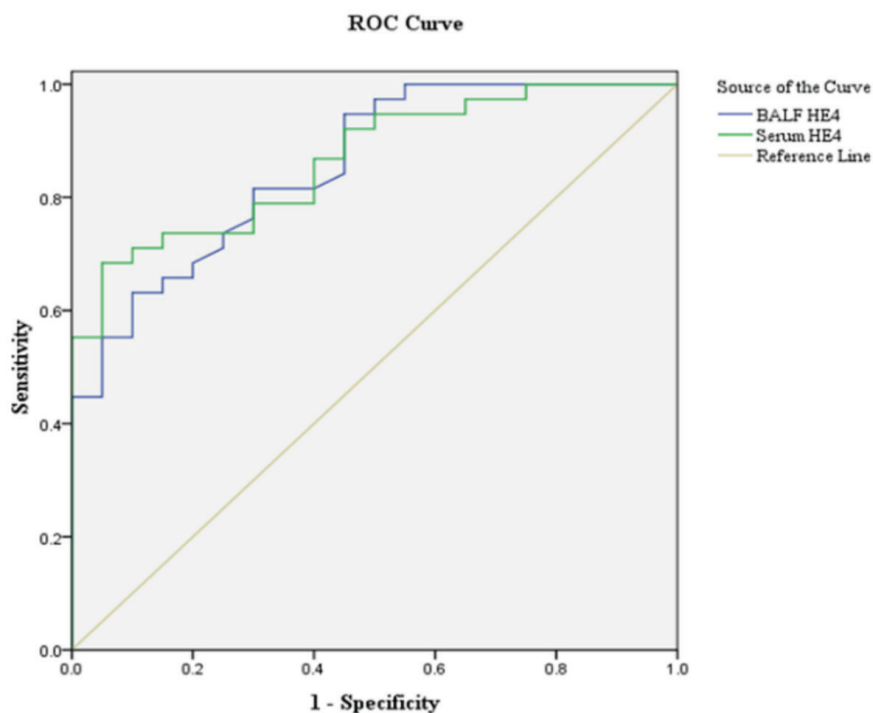


Fig. 5. Predictive capacity of serum HE4 and BALF HE4 in the presence of ILD in RA. HE4: human epididymis protein 4; BALF: bronchoalveolar lavage fluid; RA-ILD: rheumatoid arthritis-related interstitial lung disease.

Table III. Risk factors for ILD in patients with RA assessed by logistic regression analysis.

variable	Odds ratio	95% confidence interval	p-value
Univariate analysis			
Age (years)	1.075	1.021-1.131	0.006
Male	1.924	0.857-4.322	0.113
Smoking history	1.304	0.534-3.185	0.561
BMI (kg/m ²)	1.069	0.857-1.334	0.552
RA disease duration (years)	1.011	0.967-1.056	0.628
DAS28 (units)	1.208	0.872-1.672	0.256
RF titre (IU/ml)	1.001	1.000-1.002	0.118
Anti-CCP antibody titre (IU/ml)	1.001	0.995-1.006	0.846
ESR (mm/h)	1.012	0.995-1.028	0.168
CRP (mg/dl)	1.006	0.996-1.017	0.244
Serum HE4 \geq 111.1	11.847	4.242-33.086	0.000
BALF HE4 \geq 595	16.923	5.527-131.152	0.000
Multivariate analysis			
Age (years)	1.054	0.975-1.139	0.186
Serum HE4 \geq 111.1	5.082	0.986-26.188	0.052
BALF HE4 \geq 595	8.090	1.317-49.682	0.024

BMI: body mass index; DAS: disease activity score; RF: rheumatoid factor; anti-CCP: anti-cyclic citrullinated peptide; ESR: erythrocyte sedimentation rate; CRP: c-reactive protein; BALF: bronchoalveolar lavage fluid; HE4: human epididymis protein 4.

subset of pulmonary epithelial-derived tumour cell lines (29). Within these tissues, HE4 is most readily detected in the excretory ducts of both the major and minor glands. The expression of HE4 at these sites suggests that it may be released into the secretions of the glands and the epithelium. HE4 has recently been identified as being a component of

bronchoalveolar lavage (30). The exact function of HE4 has not been clearly illustrated; however, the protein has been suggested to function in the innate immune system in the lung (31, 32). It has antimicrobial and proteinase inhibitor activities and displays multiple biological activities in immunomodulation, anti-infection and fibrosis formation.

Fibroblast-derived HE4 was identified as being a mediator in kidney fibrosis in human and mouse models, which indicated a functional role of HE4 in the formation of fibrosis (14). Significant increases in the expression and localisation of HE4 were observed in patients with cystic fibrosis. In addition, HE4 showed a positive correlation with the degree of lung impairment.

Our results, regarding the association between HE4 and ILD in RA patients, were similar to the findings by Zhang *et al.*, who observed that the HE4 levels in serum and BALF were significantly elevated in patients with SSc who had ILD (20). In the present study, we also determined the diagnostic power of HE4 via ROC curve analyses. The AUC for serum HE4 was 0.782 (95% CI: 0.683–0.881) at the 111.1 pmol/L cut-off value, and the AUC for BALF HE4 was 0.856 (95% CI: 0.759–0.952) at the 595 pmol/l cut-off value for discriminating RA-ILD from RA, with appropriate qualities of sensitivity and specificity. Moreover, the multivariate logistic regression analysis demonstrated that the high levels of BALF HE4 (\geq 595 pmol/l) (odds ratio (OR)= 8.090; 95% CI=1.317–49.682; $p=0.024$) were significantly associated with RA-ILD.

Some studies have demonstrated that serum HE4 plays the role of a disease progression biomarker for some diseases, including ovarian cancer, chronic renal disease, pulmonary tuberculosis, cystic fibrosis and SSc-ILD (13, 15, 19, 33, 34). In our study, the serum and BALF HE4 concentrations showed a significant inverse correlation with DLCO% pred in RA-ILD patients. This result suggested that HE4 could be a potential biomarker for ILD and disease progression in RA patients. However, the specific mechanism of elevated HE4 levels in RA patients is unknown. We speculate that HE4 inhibited the activity of multiple matrix metalloproteinases, specifically inhibited their contribution to degrade collagen I and promoted fibrosis formation, which may be similar to the mechanics in SSc-ILD patients. However, further research is needed to confirm this hypothesis.

Until now, there has been no consensus on the definition of ILD in RA patients,

the sensitivity of the detection method, and the temporality of screening (*i.e.* RA duration at the time of screening). Clinically significant ILD will develop in approximately 10% of RA patients (35). However, preclinical ILD may be detected in up to 60% of RA patients with HRCT (4, 5, 36). The screening method based on clinical signs and symptoms will be less sensitive for detecting ILD than chest HRCT. Consequently, the ILD incidence has been increasing in recent years (37). In a recent US study involving a health insurance database, the mean total 5-year costs were \$173,405 per patient (SD: \$158,837), which suggested that the health care use and costs were substantial (8). Therefore, when considering cost-effectiveness and radiation hazard, serum biomarkers can be easily and quickly measured, which may be a priority.

Several studies investigating serum Krebs von den Lungen-6 (KL-6), surfactant protein D (SP-D) and second-generation anticyclic citrullinated peptide antibodies (anti-CCP2) have been useful biomarkers for detecting ILD in RA patients. KL-6 is a mucin-like glycoprotein expressed on the surface membrane of alveolar and bronchiolar epithelial cells that has profibrotic and anti-apoptotic effects on lung fibroblasts (38). Several studies investigating the clinical significance of KL-6 in RA patients have suggested that KL-6 serum levels are useful for detecting ILD and for evaluating disease activity. High levels of KL-6 were found to be correlated with low FVC and a high mortality rate (39). Surfactant protein D (SP-D), which is an alveolar epithelial marker, has been associated with preclinical and clinical RA-ILD (40). Anti-CCP2 antibodies could constitute an independent factor associated both with the presence and the severity of RA-ILD (41). Interestingly, in a recent study, the tumour markers CA125, CA19-9 and CEA were increased in RA-ILD and may play a role in occurrence and development, which may be diagnostic indicators for RA-ILD. Our study confirmed that serum and BALF HE4 levels were significantly increased in RA-ILD patients; a further analysis

found that HE4 had a positive correlation with HRCT fibrosis scores, which meant that a high HE4 level was an important discriminating factor of ILD, and that it may be a useful predictor for the severity of ILD in RA patients. One of any biomarkers alone may not be sufficient to completely and correctly screen/diagnose ILD in RA patients. Different biomarkers have various limitations. We speculated that a combination of biomarker assays may be more useful in clinical practice. Future studies based on a combination of biomarkers are needed to verify our preliminary findings.

Our study also had several limitations. First, it was a single-centre cross-sectional study. Therefore, it was unable to demonstrate the causality of HE for the development of ILD in RA. Further studies using longitudinal data are required to establish exact and definite causal relationships. Second, the number of patients included in this study was relatively small. Third, this study lacked a validation cohort to confirm the results. In addition, the external validity of the results in other racial cohorts would also be an important aspect to validate. Finally, the lack of histopathological examination may also limit the value of the study. Direct serum/BALF-pathology correlations will be more convincing for determining HE4 as being a biomarker for RA-ILD.

Conclusions

In conclusion, this study aimed to identify new biomarkers in RA-ILD patients, which could facilitate the early diagnosis and treatment of this disease. The results showed that HE4 was markedly elevated in both the serum and BALF of RA-ILD patients. Moreover, there was a strong correlation of HE4 with ILD disease severity in RA patients. Therefore, HE4 could be a clinically useful biomarker in screening and evaluating RA-ILD patients.

Acknowledgements

We thank all the participants who agreed to enrol in the study and are grateful to the staff of the Lung Function Lab and Bronchoscope Lab at Yantai Yuhuangding Hospital for their assistance.

References

1. MCINNES IB, SCHETT G: The pathogenesis of rheumatoid arthritis. *N Engl J Med* 2011; 365: 2205-19. <https://doi.org/10.1056/nejmra1004965>
2. FIGUS FA, PIGAM, AZZOLINI I *et al.*: Rheumatoid arthritis: extra-articular manifestations and comorbidities. *Autoimmun Rev* 2021; 20: 102776. <https://doi.org/10.1016/j.autrev.2021.102776>
3. KODURI G, NORTON S, YOUNG A *et al.*: Interstitial lung disease has a poor prognosis in rheumatoid arthritis: results from an inception cohort. *Rheumatology (Oxford)* 2010; 49: 1483-89. <https://doi.org/10.1093/rheumatology/keq035>
4. BONGARTZ T, NANNINI C, MEDINA-VELASQUEZ YF *et al.*: Incidence and mortality of interstitial lung disease in rheumatoid arthritis: a population-based study. *Arthritis Rheum* 2010; 62: 1583-91. <https://doi.org/10.1002/art.27405>
5. KELLY CA, SARAVANAN V, NISAR M *et al.*: Rheumatoid arthritis-related interstitial lung disease: associations, prognostic factors and physiological and radiological characteristics—a large multicentre UK study. *Rheumatology (Oxford)* 2014, 53: 1676-82. <https://doi.org/10.1093/rheumatology/keu165>
6. GABBAY E, TARALA R, WILL R *et al.*: Interstitial lung disease in recent onset rheumatoid arthritis. *Am J Respir Crit Care Med* 1997; 156: 528-35. <https://doi.org/10.1164/ajrccm.156.2.9609016>
7. FAZELI MS, KHAYCHUK V, WITTSTOCK K *et al.*: Rheumatoid arthritis-associated interstitial lung disease: epidemiology, risk/prognostic factors, and treatment landscape. *Clin Exp Rheumatol* 2021; 39: 1108-18. <https://doi.org/10.55563/clinexprheumatol/h9tc57>
8. RAIMUNDO K, SOLOMONO JJ, OLSON AL *et al.*: Rheumatoid arthritis-interstitial lung disease in the United States: prevalence, incidence, and healthcare costs and mortality. *J Rheumatol* 2019; 46: 360-69. <https://doi.org/10.3899/jrheum.171315>
9. HELLSTRÖM I, RAYCRAFT J, HAYDEN-LED-BETTER M *et al.*: The HE4 (WFDC2) protein is a biomarker for ovarian carcinoma. *Cancer Res* 2003; 63: 3695-3700.
10. IWAHORI K, SUZUKI H, KISHI Y *et al.*: Serum HE4 as a diagnostic and prognostic marker for lung cancer. *Tumour Biol* 2012; 33: 1141-9. <https://doi.org/10.1007/s13277-012-0356-9>
11. NAGY B, BHATTOA HP, STEIBER Z *et al.*: Serum human epididymis protein 4 (HE4) as a tumor marker in men with lung cancer. *Clin Chem Lab Med* 2014; 52: 1639-48. <https://doi.org/10.1515/cclm-2014-0041>
12. BIGNOTTI E, RAGNOLI M, ZANOTTI L *et al.*: Diagnostic and prognostic impact of serum HE4 detection in endometrial carcinoma patients. *Br J Cancer* 2011; 104: 1418-25. <https://doi.org/10.1038/bjc.2011.109>
13. LI L, WAN J, CAI G *et al.*: Value of serum human epididymis secretory protein 4 as a marker for differential diagnosis of malignant and benign gynecological diseases of patients in southern China. *Clin Chim Acta* 2016; 459: 170-6. <https://doi.org/10.1016/j.cca.2016.06.010>

14. LEBLEU VS, TENG Y, O'CONNELL JT *et al.*: Identification of human epididymis protein-4 as a fibroblast-derived mediator of fibrosis. *Nat Med* 2013; 19: 227-31. <https://doi.org/10.1038/nm.2989>
15. NAGY B, NAGY B, FILA L, CLARKE LA *et al.*: Human epididymis protein 4: a novel serum inflammatory biomarker in cystic fibrosis. *Chest* 2016; 150: 661-72. <https://doi.org/10.1016/j.chest.2016.04.006>
16. LUO J, WANG F, WAN J *et al.*: Serum human epididymis secretory protein 4 as a potential biomarker of renal fibrosis in kidney transplantation recipients. *Clin Chim Acta* 2018; 483: 216-21. <https://doi.org/10.1016/j.cca.2018.05.006>
17. RAGHU G, RICHELDI L, JAGERSCHMIDT A *et al.*: Idiopathic pulmonary fibrosis: prospective, case-controlled study of natural history and circulating biomarkers. *Chest* 2018; 154: 1359-70. <https://doi.org/10.1016/j.chest.2018.08.1083>
18. NISHIYAMA N, MASUO M, NUKUI Y *et al.*: Human epididymis protein 4 is a new biomarker to predict the prognosis of progressive fibrosing interstitial lung disease. *Respir Investig* 2021; 59: 90-8. <https://doi.org/10.1016/j.resinv.2020.08.002>
19. ZHANG M, ZHANG L, LINNING E *et al.*: Increased levels of HE4 (WFDC2) in systemic sclerosis: a novel biomarker reflecting interstitial lung disease severity? *Ther Adv Chronic Dis* 2020; 11: 2040622320956420. <https://doi.org/10.1177/2040622320956420>
20. ALETAHA D, NEOGIT, SILMAN AJ *et al.*: 2010 rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Ann Rheum Dis* 2010; 69: 1580-8. <https://doi.org/10.1136/ard.2010.138461>
21. GRAHAM BL, STEENBRUGGEN I, MILLER MR *et al.*: Standardization of Spirometry 2019 Update. An Official American Thoracic Society and European Respiratory Society Technical Statement. *Am J Respir Crit Care Med* 2019; 200: e70-e88. <https://doi.org/10.1164/rccm.201908-1590st>
22. TASHKIN DP, ELASHOFF R, CLEMENTS PJ *et al.*: Cyclophosphamide versus placebo in scleroderma lung disease. *N Engl J Med* 2006; 354: 2655-66. <https://doi.org/10.1056/nejmoa055120>
23. KAZEROONI EA, MARTINEZ FJ, FLINT A *et al.*: Thin-section CT obtained at 10-mm increments versus limited three-level thin-section CT for idiopathic pulmonary fibrosis: correlation with pathologic scoring. *Am J Roentgenol* 1997; 169: 977-983. <https://doi.org/10.2214/ajr.169.4.9308447>
24. MEYER KC, RAGHU G, BAUGHMAN RP *et al.*: An official American Thoracic Society clinical practice guideline: the clinical utility of bronchoalveolar lavage cellular analysis in interstitial lung disease. *Am J Respir Crit Care Med* 2012; 185: 1004-14. <https://doi.org/10.1164/rccm.201202-0320st>
25. KIRCHHOFF C, HABBEN I, IVELL R *et al.*: A major human epididymis-specific cDNA encodes a protein with sequence homology to extracellular proteinase inhibitors. *Biol Reprod* 1991; 45: 350-7. <https://doi.org/10.1095/biolreprod45.2.350>
26. LEBLEU VS, TENG Y, O'CONNELL JT *et al.*: Identification of human epididymis protein-4 as a fibroblast-derived mediator of fibrosis. *Nature Medicine* 2013; 19: 227-31. <https://doi.org/10.1038/nm.2989>
27. PLEBANI M, GROUP HS: HE4 in gynecological cancers: report of a European investigators and experts meeting. *Clin Chem Lab Med* 2012; 50: 2127-36. <https://doi.org/10.1515/cclm-2012-0373>
28. MONTAGNANA M, DANESE E, GIUDICI S *et al.*: HE4 in ovarian cancer: from discovery to clinical application. *Adv Clin Chem* 2011; 55: 1-20.
29. BINGLE L, SINGLETON V, BLINGLE 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; 21: 2768-73. <https://doi.org/10.1038/sj.onc.1205363>
30. WU J, KOBAYASHI M, SOUSA FA *et al.*: Differential proteomic analysis of bronchoalveolar lavage fluid in asthmatics following segmental antigen challenge. *Mol Cell Proteomics* 2005; 4: 1251-64. <https://doi.org/10.1074/mcp.m500041-mcp200>
31. BINGLE CD, VYAKARNAM A: Novel innate immune functions of the whey acidic protein family. *Trends Immunol* 2008; 29: 444-53. <https://doi.org/10.1016/j.it.2008.07.001>
32. BINGLE L, CROSS SS, HIGHAS *et al.*: WFDC2 (HE4): a potential role in the innate immunity of the oral cavity and respiratory tract and the development of adenocarcinomas of the lung. *Respir Res* 2006; 7: 61. <https://doi.org/10.1186/1465-9921-7-61>
33. WAN J, WANG Y, CAI G *et al.*: Elevated serum concentrations of HE4 as a novel biomarker of disease severity and renal fibrosis in kidney disease. *Oncotarget* 2016; 7: 67748-59. <https://doi.org/10.18632/oncotarget.11682>
34. LI L, YAO Y, LIANG J *et al.*: Serum human epididymis protein 4 concentrations are associated with severity of patients with pulmonary tuberculosis. *Clin Chim Acta* 2020; 502: 255-60. <https://doi.org/10.1016/j.cca.2019.11.009>
35. OLSON AL, SWIGRIS JJ, SPRUNGER DB *et al.*: Rheumatoid arthritis-interstitial lung disease-associated mortality. *Am J Respir Crit Care Med* 2011; 183: 372-8. <https://doi.org/10.1164/rccm.201004-0622oc>
36. MYASOEDOVA E, CROWSON CS, TURESSON C *et al.*: Incidence of extraarticular rheumatoid arthritis in Olmsted County, Minnesota, in 1995-2007 versus 1985-1994: a population-based study. *J Rheumatol* 2011; 38: 983-9. <https://doi.org/10.3899/jrheum.101133>
37. BARTELS CM, BELL CL, SHINKI K *et al.*: Changing trends in serious extra-articular manifestations of rheumatoid arthritis among United States veterans over 20 years. *Rheumatology (Oxford)* 2010; 49: 1670-5. <https://doi.org/10.1093/rheumatology/keq135>
38. KOHNO N, AWAYA Y, OYAMA T *et al.*: KL-6, a mucin-like glycoprotein, in bronchoalveolar lavage fluid from patients with interstitial lung disease. *Am Rev Respir Dis* 1993; 148:637-42. <https://doi.org/10.1164/ajrccm/148.3.637>
39. KIM HC, CHOI KH, JACOB J *et al.*: Prognostic role of blood KL-6 in rheumatoid arthritis-associated interstitial lung disease. *PLOS One* 2020; 15: e0229997. <https://doi.org/10.1371/journal.pone.0229997>
40. DOYLE TJ, PATELAS, HATABU H *et al.*: Detection of rheumatoid arthritis-interstitial lung disease is enhanced by serum biomarkers. *Am J Respir Crit Care Med* 2015; 191: 1403-12. <https://doi.org/10.1164/rccm.201411-1950oc>
41. ROCHA-MUÑOZ AD, PONCE-GUARNEROS M, GAMEZ-NAVA JI *et al.*: Anti-cyclic citrullinated peptide antibodies and severity of interstitial lung disease in women with rheumatoid arthritis. *J Immunol Res* 2015; 2015: 151626. <https://doi.org/10.1155/2015/151626>