Official Journal of the Asian Pacific Society of Respirology
Respirology
Editor's Choice

ORIGINAL ARTICLE

Serum YKL-40 is a reliable biomarker for pulmonary alveolar proteinosis

FRANCESCO BONELLA,[1] XIAOPING LONG,[1,2] XUAN HE,[1] SHINICHIRO OHSHIMO,[3] MATTHIAS GRIESE,[4] JOSUNE GUZMAN[5] AND ULRICH COSTABEL[1]

[1]*Interstitial and Rare Lung Disease Unit, Ruhrlandklinik, University of Duisburg-Essen, Essen,* [4]*Dr. von Haunersches Kinderspital, University of Munich, Munich,* [5]*General and Experimental Pathology, Ruhr University, Bochum, Germany,* [2]*Department of Respiratory Medicine, The First Affiliated Hospital of University of South China, Hunan, China and* [3]*Department of Molecular and Internal Medicine, Graduate School of Biomedical Sciences, Hiroshima University, Hiroshima, Japan*

ABSTRACT

Background and objective: **Pulmonary alveolar proteinosis (PAP) is a rare disease characterized by alveolar filling. YKL-40, a chitinase-like protein produced by macrophages and epithelial cells, is increased in patients with interstitial lung diseases. We aimed to evaluate the role of YKL-40 as a biomarker for PAP.**

Methods: **A total of 34 patients with autoimmune PAP and 50 healthy controls were studied. YKL-40 was measured by ELISA in serum and bronchoalveolar lavage fluid (BALF). Chitinase coding gene polymorphisms (CHI3L1−329 and −131) were detected by PCR and pyrosequencing. Correlations between serum YKL-40 levels and disease outcome were analysed.**

Results: **Baseline serum and BALF levels of YKL-40 were higher in PAP patients than in controls (286 ± 27 ng/mL vs 42 ± 4 ng/mL, $P < 0.0001$; 323 ± 36 ng/mL vs 3 ± 1 ng/mL, $P < 0.0001$, respectively). Serum YKL-40 levels correlated with diffusing capacity of the lung for carbon monoxide (DL_{CO}) at baseline ($P = 0.002$) and over time ($P < 0.0001$). Patients with disease progression had higher baseline serum YKL-40 levels than those who remained stable or improved ($P < 0.0001$). A baseline cut-off level of 300 ng/mL was predictive of disease progression (HR (hazard ratio): 7.875, $P = 0.001$). The presence of the G allele was associated with higher serum and BALF levels of YKL-40.**

Conclusion: **YKL-40 is elevated in serum and BALF of PAP patients, and may be of clinical utility to predict outcome in PAP.**

Key words: biomarker, disease outcome, pulmonary alveolar proteinosis, YKL-40.

SUMMARY AT A GLANCE

We evaluated the role of YKL-40, a chitinase-like protein produced by macrophages and epithelial cells, as a biomarker to assess disease activity and outcome in pulmonary alveolar proteinosis (PAP). YKL-40 was elevated in serum and bronchoalveolar lavage fluid (BALF) of PAP patients, and correlated with respiratory impairment and disease outcome.

Abbreviations: $AaDO_2$, alveolar-arterial oxygen gradient; AUC, area under the curve; BALF, bronchoalveolar lavage fluid; BMI, body mass index; CEA, carcinoembryonic antigen; CYFRA 21-1, Cytokeratinfragment 21-1; DL_{CO}, diffusing capacity of the lung for carbon monoxide; DSS, disease severity score; ELISA, enzyme-linked immunosorbent assay; FEV_1, forced expiratory volume in 1 s; FVC, forced vital capacity; GM-CSF, granulocyte macrophage colony-stimulating factor; HR, hazard ratio; HRCT, high-resolution computed tomography; Ig, immunoglobulin; ILD, interstitial lung disease; IRB, Institutional Review Board; KL-6, Krebs von den Lungen-6; LD, linkage disequilibrium; LDH, lactate dehydrogenase; PaO_2, partial pressure of oxygen in arterial blood; PAP, pulmonary alveolar proteinosis; PCR, polymerase chain reaction; ROC, receiver operating characteristic; SaO_2, arterial oxygen saturation; SNP, single nucleotide polymorphism; TLC, total lung capacity; WLL, whole-lung lavage.

INTRODUCTION

Pulmonary alveolar proteinosis (PAP) is a rare disorder with progressive intra-alveolar filling by surfactant lipoproteins.[1] Primary autoimmune PAP is the most common form and is characterized by the occurrence of serum granulocyte macrophage colony-stimulating factor (GM-CSF) autoantibodies in 100% of cases.[2–6] Causes of secondary PAP include several underlying conditions, mostly haematological disorders, infections or dust and fume inhalation.[2–6] Many patients reach remission after one whole-lung lavage (WLL),[2–7] which remains the standard treatment, but there is a subgroup of patients requiring repeated WLLs and eventually GM-CSF therapy.[7] One of the unmet needs in PAP

Correspondence: Francesco Bonella, Interstitial and Rare Lung Disease Unit, Ruhrlandklinik, University Hospital, University of Duisburg-Essen, Tünchender Weg 40, 45239 Essen, Germany. Email: francesco.bonella@ruhrlandklinik.uk-essen.de

Received 26 August 2016; invited to revise 30 November 2016; revised 1 February 2017; accepted 8 March 2017 (Associate Editor: Yuben Moodley).

is the identification of non-invasive biomarkers, which can also predict disease outcome. Recently, Krebs von den Lungen-6 (KL-6), an established biomarker for interstitial lung diseases (ILDs) especially in Japan, has also been shown to be elevated in PAP and of potential clinical value in PAP patients.[8]

YKL-40 is a glycoprotein belonging to the chitinase family and is mainly produced by neutrophils, macrophages, synovial and malignant cells. In acute and chronic inflammatory diseases with high remodelling of the extracellular matrix such as systemic sclerosis, rheumatoid arthritis, inflammatory bowel disease and asthma, YKL-40 has been found to be increased.[9] Recently, YKL-40 has been proposed as a disease marker for ILDs and sarcoidosis.[10–14] Single nucleotide polymorphisms (SNPs) in the gene encoding for YKL-40 (CHI3L1) have been associated with increased susceptibility to asthma and with the extent of liver fibrosis.[15,16] Serum YKL-40 levels seem to be influenced by CHI3L1 SNPs,[10,11,13,16,17] which is of importance in the clinical application of this biomarker.

In the present study, we investigated the potential clinical utility of YKL-40 as a biomarker for PAP. As PAP is characterized by a certain degree of inflammation and interstitial remodelling, we aimed to verify whether serum and BALF (bronchoalveolar lavage fluid) YKL-40 concentrations are elevated and correlate with disease outcome in PAP. Part of the study has been previously reported as abstract.[18]

METHODS

Study subjects

The study was retrospective and included 34 consecutive patients with autoimmune PAP investigated in the Ruhrlandklinik from January 2007 to December 2013 and 50 healthy controls. PAP was diagnosed taking into account characteristic HRCT and BALF findings, and/or histology when available.[4,19] GM-CSF autoantibodies were positive in all patients. Approval by the local Institutional Review Board (IRB) (number 10-4397) and written informed consent from patients and healthy controls were obtained.

Definition of disease progression

The criteria used to define disease progression were worsening of self-reported symptoms (dyspnoea, cough, chest pain and weight loss), and/or of lung function (decrease in FVC (forced vital capacity) > 10% and/or DL_{CO} (diffusing capacity of the lung for carbon monoxide) > 15%) and/or of chest imaging (increase in existing or demonstration of new infiltrates characteristic of PAP) since the last follow-up visit.[20,21] Otherwise, the patients were considered as stable or improved.

Measurement of YKL-40, GM-CSF antibodies and other serum biomarkers

Serum samples were collected at the time of first evaluation in all patients. In 19 patients, further samples (range: 2-19 samples per patient) were collected during follow-up, every 7 ± 1 months on average. All samples were stored at either −20 or −80°C until analysis. BALF was performed via a fibre-optic bronchoscope as previously described[22] and supernatant was stored at −80°C. YKL-40 in serum and BALF were measured by ELISA (Quidel, San Diego, CA, USA) as described previously[13] whilst GM-CSF antibody levels were quantified by a modified ELISA using a human IgG1 antibody specific for human GM-CSF (BI01049904; Boehringer Ingelheim, Ingelheim am Rhein, Germany).[23–25] The detection limit of this GM-CSF assay was 0.2 μg/mL, and 9 μg/mL was considered the upper limit of normal.[26] Carcinoembryonic antigen (CEA) and lactate dehydrogenase (LDH) were measured in serum by automatized spectrophotometry (Konelab T series; Thermo Fisher Scientific, Vantaa, Finland), with a normal value for CEA of <2.5 ng/mL and for LDH of <225 U/L.

CHI3L1 SNPs detection and genotyping

Genomic DNA extraction from peripheral blood leucocytes was made via silica-membrane-based nucleic acid purification Kit (Qiagen DNA Mini Kit; Qiagen, Germantown, MD, USA). DNA aliquots were stored at −80°C until use. We selected two SNPs (rs10399931 for −329G/A and rs4950928 for 131G/C) in the promoter region in CHI3L1 gene (chromosome 1q32.1). For the two SNPs, we used the following primers: (i) for CHI3L1 −329, 5′-GAGGAAGGCTGGGAAAT GC-3′ (forward), 3′-GCCACTAGGGTGATGATGGG-5′ (reverse) and 5′-CGGCTGAGTCACATCTC-3′ (sequencing); (ii) for CHI3L1 −131, 5′-CCCATAAAAGGGCTGGTTTG-3′ (forward), 3′-TTCCCCAGGCCCTGTACTTC-5′ (reverse) and 5′-GTCCCACTCCACTCCC-3′ (sequencing). Target sequence amplification was done by PCR as previously described.[11,27] Purified PCR products (Qiagen, Cat. No. 51104) were sequenced with ABI Prism 3700 DNA Analyser (PyroMark Q24; Qiagen) to determine the genotypes.

Pulmonary function tests

Lung function was measured at the same day as blood sampling. Tests included FVC, forced expiratory volume in 1 s (FEV_1), total lung capacity (TLC), DL_{CO}, partial pressure of oxygen in arterial blood (PaO_2), arterial oxygen saturation (SaO_2) and alveolar-arterial oxygen gradient ($AaDO_2$).[28]

Statistical analysis

Normal distribution of continuous variables was verified using the Kolmogorov–Smirnov test. Parametric data are expressed as mean ± SE while categorical variables as either a percentage of the total or numerically. Student's t-test or Wilcoxon's rank test were used to compare continuous variables in two independent groups. Categorical variables were compared by using chi-square or Fischer's exact test. Spearman's correlation coefficient was obtained for bivariate correlations. Changes of lung function parameters and biomarkers over time were corrected for the interval time (months). We used receiver operating characteristic (ROC) curve analysis to test the association of serum and BALF YKL-40 levels at baseline with disease outcome. Haploview 4.0 software (Broad Institute, MIT, Cambridge, MA, USA) was applied to calculate linkage

disequilibrium (LD, D coefficient) between rs10399931 (CHI3L1 –329) and rs4950928 (CHI3L1 –131).[29] Cox's proportional hazard regression model was used for uni- and multivariate analyses of prognostic factors. The Kaplan–Meier method with log-rank test was used to investigate the association between the identified serum YKL-40 cut-off levels with the disease outcome. Significance was set for *P*-values of <0.05. All statistical analyses were performed using SPSS 17.0 (SPSS Inc., Chicago, IL, USA).

RESULTS

Baseline characteristics

Demographics, patients' characteristic and disease outcome are reported in Table 1. The median follow-up time was 543 days (range: 90–1281). A total of 21 patients had already received at least one WLL before the first evaluation. During follow-up, we treated all patients with disease progression ($n = 17$) with WLL. A subgroup of these patients ($n = 11$) needed repeated WLL prior to reaching disease stabilization.

YKL-40 in serum and BALF

Serum YKL-40 levels were significantly higher in PAP patients than in healthy controls (286 ± 27 ng/mL vs 42 ± 4 ng/mL, $P < 0.0001$), the same was seen for YKL-40 in BALF (323 ± 36 ng/mL vs 3 ± 1 ng/mL, $P < 0.0001$) (Fig. 1). Serum YKL-40 levels were higher in patients with a smoking history than in non-smokers (322 ± 30 ng/mL vs 172 ± 42 ng/mL, $P = 0.020$) (Fig. S1A (Supplementary Information)). Smoking habits did not affect serum YKL-40 levels in healthy subjects (Fig. S1B (Supplementary Information)). Serum YKL-40 levels at baseline were higher in patients who had disease progression during follow-up than in those who remained stable (376 ± 37 ng/mL vs 187 ± 21 ng/mL, $P < 0.0001$) (Fig. S2 (Supplementary Information)). Change of serum YKL-40 levels from baseline over time was significantly greater in those patients who experienced disease progression compared with those who remained stable (+237 ± 136% vs –3.3 ± 22% from baseline, respectively, $P = 0.042$).

Diagnostic value of baseline serum and BALF YKL-40 levels for PAP

ROC analysis showed, at a cut-off level of 99 ng/mL for serum YKL-40 levels, a sensitivity of 94%, specificity of 99.6% and an accuracy of 97% to discriminate PAP from healthy controls (AUC (area under the curve) = 0.988, 95% CI: 0.972–1.003, $P < 0.0001$); at a cut-off level of 20 ng/mL, BALF YKL-40 levels showed

Table 1 Baseline demographics and characteristics of all studied subjects and of PAP subgroups according to disease outcome (all data collected at baseline)

	All subjects			PAP patients		
Variables	PAP $n = 34$	Controls[†] $n = 50$	*P* value	Stable/improved $n = 17$	Progressed $n = 17$	*P* value
Gender (m/f), *n*	19/15	37/13	n.s.[‡]	7/10	12/5	n.s.[‡]
Age (years)	49 ± 2	42 ± 2	n.s.	52 ± 2	46 ± 4	n.s.
Smoking habits (non/ex/current), *n*	5/12/17	33/7/10	<0.001	4/4/9	1/8/8	n.s.[‡]
DSS (grade)[§]	2.7 ± 0.2	n.a.	n.a.	2.4 ± 0.2	3.1 ± 0.3	0.032
BMI (kg/m^2)	26 ± 1	24 ± 1	n.s.	27 ± 1	25 ± 1	n.s.
PaO_2 (mm Hg)	72 ± 3	86 ± 4	0.003	79 ± 3	64 ± 3	0.002
$AaDO_2$ (mm Hg)	37 ± 2	14 ± 2	<0.001	30 ± 3	43 ± 3	0.006
FVC (% predicted)	82 ± 3	90 ± 3	0.005	89 ± 3	74 ± 3	0.002
FEV_1 (% predicted)	77 ± 2	88 ± 9	0.012	80 ± 4	73 ± 3	n.s.
TLC (% predicted)	80 ± 3	89 ± 3	n.s.	87 ± 4	72 ± 4	0.010
DL_{CO} (% predicted)	57 ± 4	85 ± 4	<0.001	67 ± 4	44 ± 4	0.001
GM-CSF autoantibody (μg/mL)[¶]	50 ± 4	n.a.	n.a.	55 ± 7	46 ± 6	n.s.
Serum YKL-40 (ng/mL)	286 ± 27	39 ± 4	<0.001	187 ± 21	376 ± 37	<0.001
BALF YKL-40 (ng/mL)	314 ± 36	3 ± 1	<0.001	198 ± 21	353 ± 28	<0.001
Serum LDH (U/L)[¶]	284 ± 16	151 ± 11	<0.001	223 ± 13	338 ± 20	<0.001
Serum CEA (ng/mL)[¶]	8.5 ± 2	1.6 ± 0.1	<0.001	6 ± 1	13 ± 2	0.035

Unless otherwise indicated, values are expressed as mean ± SE (range).

[†]Lung function variables were available in 18 healthy controls.

[‡]Fisher's exact test, for all other comparisons Student's t-test was used.

[§]DSS 1 (no symptoms and $PaO_2 \geq 70$ mm Hg), DSS 2 (symptomatic and $PaO_2 \geq 70$ mm Hg), DSS 3 (60 mm Hg $\leq PaO_2 <$ 70 mm Hg), DSS 4 (50 mm Hg $\leq PaO_2 <$ 60 mm Hg) and DSS 5 ($PaO_2 <$ 50 mm Hg).

[¶]The cut-off of normality for each biomarker is reported in *Methods* section.

$AaDO_2$, alveolar-arterial oxygen gradient; BALF, bronchoalveolar lavage fluid; CEA, carcinoembryonic antigen; DL_{CO}, diffusing capacity of the lung for carbon monoxide; DSS, disease severity score; FEV_1, forced expiratory volume in 1 s; FVC, forced vital capacity; GM-CSF, granulocyte macrophage colony-stimulating factor; LDH, lactate dehydrogenase; n.a., not available; n.s., not significant; PaO_2, partial pressure of oxygen in arterial blood; PAP, pulmonary alveolar proteinosis; SE, standard error; TLC, total lung capacity.

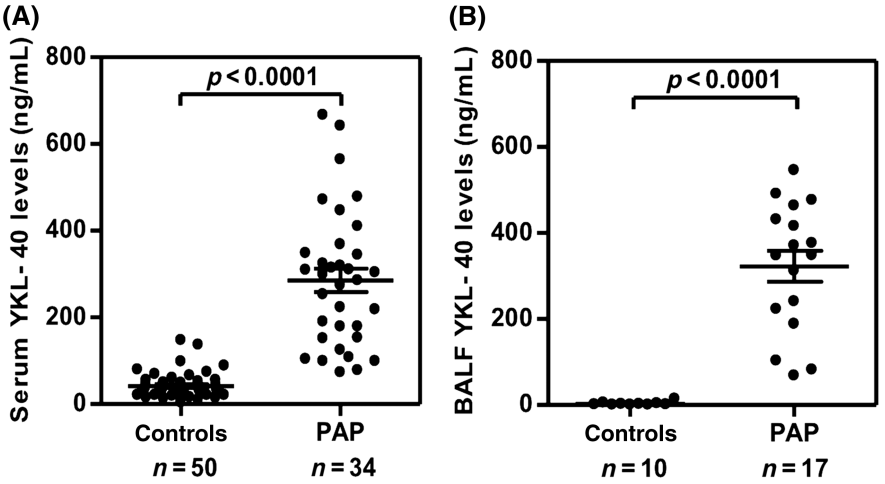

Figure 1 Serum (A) and BALF (B) YKL-40 levels of the study population at baseline. Dots represent single patients. Bars represent mean values. BALF, bronchoalveolar lavage fluid; PAP, pulmonary alveolar proteinosis.

a sensitivity of 100%, specificity of 100% and an accuracy of 100% (AUC = 1.0, $P < 0.0001$).

Correlations between serum YKL-40 and pulmonary function tests

Serum YKL-40 levels correlated inversely with DL_{CO} at baseline (r = –0.534, $P = 0.002$), PaO_2 (r = –0.427, $P = 0.012$) and positively with $AaDO_2$ (r = 0.437, $P = 0.011$) (Fig. S3A–C (Supplementary Information)). Changes in serum YKL-40 levels correlated negatively with changes in DL_{CO} over time (r = –0.554, $P < 0.0001$) (Fig. S3D (Supplementary Information)).

Other correlations

Levels of YKL-40 in serum or BALF were not affected by age, BMI, GM-CSF autoantibody or cumulative number of WLL in those who received treatment after baseline (data not shown). Baseline serum and BALF YKL-40 showed a significant correlation (r = 0.678, $P = 0.0003$) (Fig. S4A (Supplementary Information)). Serum YKL-40 also correlated with serum LDH (r = 0.388, $P = 0.023$) (Fig. S4B (Supplementary Information)) and BALF lymphocyte counts (r = 0.631, $P = 0.028$) (Fig. S4C (Supplementary Information)) at baseline. We found a significant inverse correlation between serum YKL-40 and the time to next WLL (r = –0.497, $P = 034$) as well as a weak direct correlation between the decrease of serum YKL-40 during follow-up and the time to next WLL (r = 0.26, $P = 0.026$).

Predictive role of serum YKL-40 levels at baseline for PAP outcome

Using ROC analysis, we found that 300 ng/mL was the best serum YKL-40 cut-off level (sensitivity 80% and specificity 93%, respectively) to predict disease progression (AUC = 0.895, $P < 0.0001$). At a cut-off level of 266 U/L, serum LDH levels showed similar results (sensitivity 82%, specificity 82%, AUC = 0.869, $P < 0.0001$). Serum GM-CSF autoantibody levels and disease severity score (DSS) grade had no predictive value for disease progression ($P = 0.138$ and $P = 0.055$, respectively) (Fig. 2A). According to the best cut-off value obtained in the ROC analysis, we also divided the patients into the high YKL-40 group ($n = 16$) with baseline concentrations >300 ng/mL, and the low YKL-40 group ($n = 18$) (≤300 ng/mL). Patients' characteristics according to this YKL-40 cut-off value are shown in Table 2.

Moreover, we performed univariate and multivariate analyses to investigate the association of the cut-off value with disease progression. Baseline serum YKL-40 at a cut-off level of >300 ng/mL was strongly associated with an increased risk of disease progression in the univariate analysis (HR (hazard ratio): 7.875, 95% CI: 2.122–28.043, $P = 0.001$) (Table 3). In the multivariate analysis, serum YKL-40 at a cut-off level of >300 ng/ml remained predictive for disease progression after adjustment for age, gender, smoking history, baseline PaO_2, $AaDO_2$, FVC and DL_{CO} as covariates (HR: 6.201, 95% CI: 0.866–55.704, $P = 0.021$) (Table 3).

Kaplan–Meier analysis for disease progression

The median follow-up time was 543 days (range: 90–1281). The Kaplan–Meier analysis confirmed that serum YKL-40 at 300 ng/mL was predictive for disease progression (Fig. 2B). The disease progression rate was 88% in the high YKL-40 group ($n = 16$) and 11% in the low YKL-40 group ($n = 18$) (log-rank test, $P = 0.011$).

CHI3L1 –329 and –131 SNPs and YKL-40 levels

A total of 50 healthy controls and 23 PAP patients underwent CHI3L1 –329 and –131 SNPs and genotyping investigations. The alleles of both SNPs were in Hardy–Weinberg equilibrium in both groups (data not shown). The LD between CHI3L1 –329 and –131 SNPs was high ($D' = -0.862$, $r^2 = 0.5563$). The genotypes did not correlate with gender, age or lung function tests (data not shown). Serum YKL-40 levels were higher in PAP patients with the CHI3L1 –329 G/G genotype than in those with A/G genotype (412 ± 51 ng/mL vs 225 ± 30 ng/mL, $P = 0.008$) (Fig. 3). As expected, for

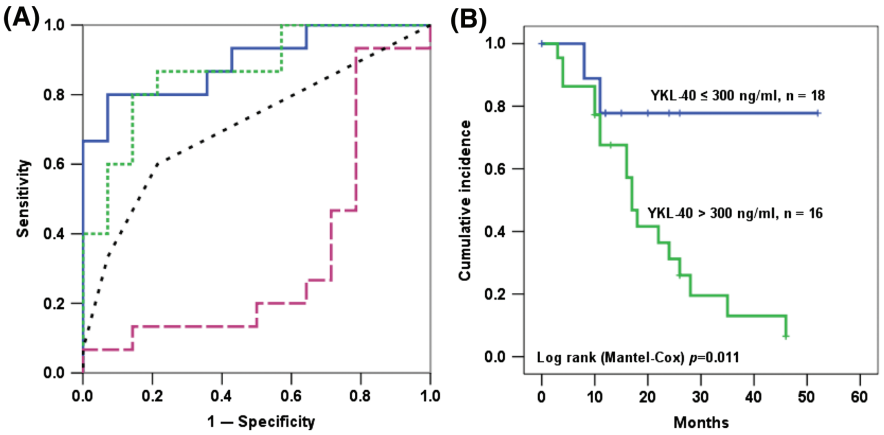

Figure 2 Predictive value of baseline serum YKL-40 levels. (A) Receiver operating characteristic (ROC) curves showing the value of baseline serum YKL-40 (——), lactate dehydrogenase (LDH, ·····), granulocyte macrophage colony-stimulating factor (GM-CSF, - - - - -) and disease severity score (DSS, ·····) for predicting disease progression. (B) Kaplan–Meier curve of serum YKL-40 at a cut-off level of 300 ng/mL as a predictive factor for disease progression.

CHI3L1 −131 SNP, similar results were seen (396 ± 44 ng/mL vs 208 ± 33 ng/mL, $P = 0.005$, for C/C and C/G genotypes, respectively). A similar association between SNP genotypes and YKL-40 levels was seen in BALF (data not shown).

DISCUSSION

In this study, we found that PAP patients had higher concentrations of YKL-40 in serum and BALF in comparison with healthy controls. The serum YKL-40 levels correlated with lung function impairment and disease outcome. Moreover, serum YKL-40 was an independent predictor of disease progression in the multivariate analysis.

YKL-40 levels have not been explored in Caucasian PAP patients before. In a Chinese PAP cohort, elevated serum levels of YKL-40 have been recently reported, comparable with those found in our patients.[30] Regarding the possible cellular source of YKL-40 in PAP, this protein can be produced by several types of cells, such as activated human macrophages, neutrophils and bronchial epithelial cells.[31,32] YKL-40 has been found to be present also in the BALF in PAP and various ILDs.[10,30] The accumulation of PAP in the alveolar space may be rather due to overproduction by the hyperplastic type II pneumocytes rather than by the functionally impaired alveolar macrophages. The elevation of YKL-40 in serum can be explained by the spill over through the air–blood barrier, which is highly permeable in the affected lungs.

Table 2 Characteristics of the patients stratified according to serum YKL-40 predictive cut-off value for disease progression ($n = 34$)

Variables	YKL-40 ≤ 300 ng/mL $n = 18$	YKL-40 > 300 ng/mL $n = 16$	*P*-value
Gender (m/f), *n*	9/9	10/6	n.s.†
Age (years)	48 ± 3	50 ± 3	n.s.
Smoking history (no/yes), *n*	4/14	1/15	n.s.†
Clinical course			
PAP-related death (yes/no), *n*	1/17	3/13	n.s.†
Disease progression (yes/no), *n*	3/15	14/2	<0.001†
DSS (grade)	2.4 ± 0.2	3.1 ± 0.3	0.049
PaO_2 (mm Hg)	77 ± 3	67 ± 4	0.047
$AaDO_2$ (mm Hg)	31 ± 3	44 ± 3	0.006
FVC (% predicted)	87 ± 4	76 ± 4	0.040
FEV_1 (% predicted)	79 ± 4	74 ± 3	n.s.
TLC (% predicted)	86 ± 4	73 ± 4	0.045
DL_{CO} (% predicted)	67 ± 4	45 ± 4	0.001
GM-CSF autoantibody (μg/mL)	57 ± 6	42 ± 6	0.080
Serum YKL-40 (ng/mL)	172 ± 17	414 ± 30	<0.001
Serum LDH (U/L)	238 ± 17	328 ± 22	0.003
Serum CEA (ng/mL)	5.6 ± 1.5	7.6 ± 2.1	n.s.

Unless otherwise indicated, values are expressed as mean ± SE.

†Chi-square or Fischer's exact test, for all other comparisons Student's t-test was used.

$AaDO_2$, alveolar-arterial oxygen gradient; CEA, carcinoembryonic antigen; DL_{CO}, diffusing capacity of the lung for carbon monoxide; DSS, disease severity score; FEV_1, forced expiratory volume in 1 s; FVC, forced vital capacity; GM-CSF, granulocyte macrophage colony-stimulating factor; LDH, lactate dehydrogenase; n.s., not significant; PaO_2, partial pressure of oxygen in arterial blood; PAP, pulmonary alveolar proteinosis; SE, standard error; TLC, total lung capacity.

Table 3 Uni- and multivariate Cox proportional analyses for the risk of disease progression

Variables	HR	95% CI		*P*-value
		Lower	Upper	
Univariate analysis				
Serum YKL-40 ng/mL (continuous)	1.006	1.003	1.009	<0.001
Serum YKL-40 > 300 ng/mL (binary)	7.875	2.122	28.043	0.001
Age, year (continuous)	0.959	0.922	1.014	0.078
Gender (female = 1)	3.642	1.517	11.469	0.027
Smoking history (positive = 1)	3.456	0.456	26.193	0.230
DSS (grade)	1.569	1.012	2.433	0.044
Serum LDH U/L (continuous)	1.013	1.007	1.019	0.001
GM-CSF autoantibody μg/mL (continuous)	0.986	0.963	1.011	0.268
PaO_2 mm Hg (continuous)	0.944	0.905	0.985	0.008
$AaDO_2$ mm Hg (continuous)	1.060	1.018	1.104	0.004
FVC % predicted (continuous)	0.966	0.935	0.998	0.040
DL_{CO} % predicted (continuous)	0.950	0.913	0.989	0.012
Multivariate analysis[†]				
Serum YKL-40 ng/mL (continuous)	1.009	1.001	1.016	0.026
Gender (female = 1)	16.054	2.170	118.750	0.007
Serum LDH U/L (continuous)	1.024	1.004	1.044	0.018
Serum YKL-40 > 300 ng/mL (binary)	6.201	0.868	55.704	0.021
Gender (female = 1)	19.802	2.177	204.188	0.120
Serum LDH U/L (continuous)	1.051	1.015	1.088	0.050

[†]Adjustment for gender, DSS, serum LDH, baseline PaO_2 (mm Hg), $AaDO_2$ (mm Hg), FVC (% predicted) and DL_{CO} (% predicted). The results were summarized as HRs, representing the relative risk to develop disease progression as a specific characteristic at baseline.

$AaDO_2$, alveolar-arterial oxygen gradient; CI, confidence interval; DL_{CO}, diffusing capacity of the lung for carbon monoxide; DSS, disease severity score; FVC, forced vital capacity; GM-CSF, granulocyte macrophage colony-stimulating factor; HR, hazard ratio; LDH, lactate dehydrogenase; PaO_2, partial pressure of oxygen in arterial blood.

It is crucial for a reliable biomarker to be not dependent on age, gender or clinical variables, otherwise there is no additional clinical value to its use. Although we found that serum and BALF levels of YKL-40 were not influenced by gender or age, the impact of smoking habits is still not clear. In our cohort, serum YKL-40 levels were higher in former and current smokers than in never smokers, as has previously been reported for adult healthy controls.[33] Smoking causes hyperplasia of type II pneumocytes and disruption of the air--blood membrane[34] and could enhance both production and spill over of YKL-40. YKL-40 has been found overexpressed in alveolar macrophages and pneumocytes in cigarette smokers and in cigarette smoke-exposed mice.[35] We think that smoking habits should be taken into consideration when interpreting YKL-40 serum levels, but the influence of smoking on the predictive value of YKL-40 for disease outcome seems not to be relevant.

Serum YKL-40 inversely correlated with DL_{CO} and blood gas analyses at baseline, and interestingly, changes in DL_{CO} correlated with serum changes in YKL-40 over time. Furthermore, the serum levels of YKL-40 correlated well with LDH, a non-specific marker of tissue injury widely used to monitor the course of acute lung disease.[36] It is known that gas exchange values (DL_{CO} and PaO_2) rather than lung function tests (FVC or TLC) are more reliable parameters for the clinical follow-up in PAP patients.[5,25,37] Therefore, YKL-40 could be of utility as a surrogate of

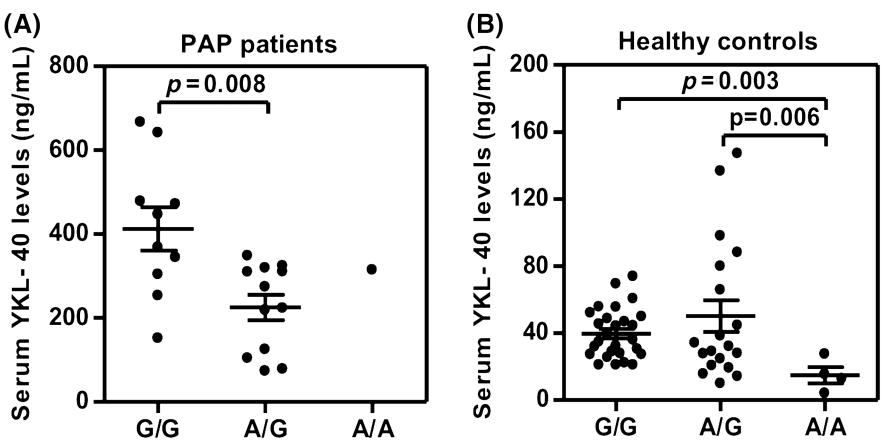

Figure 3 Serum YKL-40 levels in PAP patients (A) and healthy controls (B) according to CHI3L1 -329 genotype (results for -131 SNP are not shown since the SNPs are in high linkage disequilibrium). Dots represent single patients. Bars represent mean values.

DL_{CO}, whose measurement is burdened by fluctuations.[38]

We found that patients with progressive disease had higher levels of serum YKL-40 than those who remained stable or improved. Serum YKL-40 levels >300 ng/mL selected through ROC analysis predicted disease progression with a sensitivity of 80% and specificity of 93%, similar to serum LDH (sensitivity of 82% and specificity of 82%). However, for serum LDH, it was not possible to identify a cut-off level as independent predictor of disease outcome, as indicated by the results of multivariate analysis. Whether YKL-40 is superior to LDH as biomarker for PAP needs further investigation. The third biomarker, GM-CSF autoantibody, was not associated with disease outcome, as shown previously.[5,21] Disease progression was seen in 88% of patients with serum YKL-40 levels > 300 ng/mL but only in 11% of patients under this threshold. Moreover, in the multivariate analysis, a serum YKL-40 level >300 ng/mL, but not YKL-40 as continuous variable, was significantly associated with disease progression with a pretty high HR (6.2). Although the cut-off value identified in the present study cannot be used in other cohorts without appropriate validation, this does not reduce the potential clinical utility of YKL-40 to identify patients with higher risk of disease progression.

Moreover, we found that baseline YKL-40 levels were inversely associated with time to next WLL, a surrogate measure to assess treatment response in those patients who require multiple WLL over time.[20,39] Further investigations are needed to evaluate the role of YKL-40 to assess treatment response.

Beside clinical-physiological parameters, such as the DSS,[20] two biomarkers have been extensively studied to evaluate outcome in PAP, Cytokeratinfragment 21-1 (CYFRA 21-1)[40] and KL-6.[21,41] CYFRA 21-1 has been successfully used to evaluate the response to GM-CSF treatment[40] and KL-6 to identify patients with complicated disease course.[21,41] We were not able to compare serum YKL-40, KL-6 and CYFRA 21-1 in the present study because of the low number of patients tested for all biomarkers.

With regard to the distribution of the YKL-40 encoding gene polymorphisms, we looked at CHI3L1 SNPs 10399931 and rs4950928 as they have previously shown to explain 23% of variation in serum YKL-40 levels in healthy subjects.[11] In our cohort, PAP patients with the CHI3L1 −329 G/G and CHI3L1 −131 C/C genotype showed higher levels of serum YKL-40 than those with CHI3L1 −329 A/G and CHI3L1 −131 C/G. In healthy controls, only a tendency was seen. As the different genotypes were not independently associated with gender, age or lung function parameters, we can conclude that these SNPs are likely to explain the variation of serum YKL-40 in PAP between individuals, but they are not associated with a PAP sub-phenotype or a specific disease outcome.

Finally, the limitations of the present study need to be discussed. First, the sample size is too small to draw final conclusions, especially on the association between CHI3L1 genotype and disease severity and outcome. Second, we were unable to investigate how serum YKL-40 levels are related to functional response to WLL treatment, as we did not collect systematically lung function data after each procedure. Third, immunohistochemistry to demonstrate YKL-40 in biopsy samples as well as alveolar macrophages cell culture was not performed, therefore we can only speculate on the source and distribution of this protein in PAP. In addition, we did not perform a comparison study between YKL-40 and established biomarkers for PAP, such as KL-6 or CYFRA 21-1.

In conclusion, we show that serum YKL-40 could be a reliable biomarker for predicting disease outcome in PAP. This marker may allow a better stratification of PAP patients. A multicentre validation and a comparative biomarker study are required to determine whether serum YKL-40 could be routinely used as a biomarker in PAP patients and is of additional advantage.

Acknowledgements

This study was supported by the German Federal Ministry of Education and Research (EuPAPNet, ERARE Joint Call 2009, project number 01GM1011A) and Arbeitsgemeinschaft zur Förderung der Pneumologie an der Ruhrlandklinik (AFPR).

Disclosure Statement

F.B. has received speaker fees, advisory board honoraria or grants from InterMune International AG which became a wholly owned subsidiary of Roche in 2014, Boehringer Ingelheim, Gilead, Serendex, Centocor and F. Hoffmann-La Roche Ltd.

REFERENCES

1 Rosen SH, Castleman B, Liebow AA. Pulmonary alveolar proteinosis. *N. Engl. J. Med.* 1958; **258**: 1123–42.
2 Costabel U, Guzman J. Pulmonary alveolar proteinosis: a new autoimmune disease. *Sarcoidosis Vasc. Diffuse Lung Dis.* 2005; **22** (Suppl. 1): S67–73.
3 Kitamura T, Tanaka N, Watanabe J, Uchida K, Kanegasaki S, Yamada Y, Nakata K. Idiopathic pulmonary alveolar proteinosis as an autoimmune disease with neutralizing antibody against granulocyte/macrophage colony-stimulating factor. *J. Exp. Med.* 1999; **190**: 875–80.
4 Carey B, Trapnell BC. The molecular basis of pulmonary alveolar proteinosis. *Clin. Immunol.* 2010; **135**: 223–35.
5 Inoue Y, Trapnell BC, Tazawa R, Arai T, Takada T, Hizawa N, Kasahara Y, Tatsumi K, Hojo M, Ichiwata T *et al.*; Japanese Center of the Rare Lung Diseases Consortium. Characteristics of a large cohort of patients with autoimmune pulmonary alveolar proteinosis in Japan. *Am. J. Respir. Crit. Care Med*. 2008; **177**: 752–62.
6 Seymour JF, Presneill JJ. Pulmonary alveolar proteinosis: progress in the first 44 years. *Am. J. Respir. Crit. Care Med.* 2002; **166**: 215–35.
7 Leth S, Bendstrup E, Vestergaard H, Hilberg O. Autoimmune pulmonary alveolar proteinosis: treatment options in year 2013. *Respirology* 2013; **18**: 82–91.
8 Ishikawa N, Hattori N, Yokoyama A, Kohno N. Utility of KL-6/MUC1 in the clinical management of interstitial lung diseases. *Respir. Investig.* 2012; **50**: 3–13.
9 Johansen JS. Studies on serum YKL-40 as a biomarker in diseases with inflammation, tissue remodelling, fibroses and cancer. *Dan. Med. Bull.* 2006; **53**: 172–209.
10 Korthagen NM, van-Moorsel CHM, Zanen P, Ruven HJ, Grutters JC. Evaluation of circulating YKL-40 levels in idiopathic interstitial pneumonias. *Lung* 2014; **192**: 975–80.
11 Kruit A, Grutters JC, Ruven HJ, van-Moorsel CC, van-den-Bosch JM. A CHI3L1 gene polymorphism is associated with serum levels of YKL-40, a novel sarcoidosis marker. *Respir. Med.* 2007; **101**: 1563–71.

12 Johansen JS, Milman N, Hansen M, Garbarsch C, Price PA, Graudal N. Increased serum YKL-40 in patients with pulmonary sarcoidosis – a potential marker of disease activity? *Respir. Med.* 2005; **99**: 396–402.
13 Korthagen NM, van-Moorsel CH, Barlo NP, Ruven HJ, Kruit A, Heron M, van-den-Bosch JM, Grutters JC. Serum and BALF YKL-40 levels are predictors of survival in idiopathic pulmonary fibrosis. *Respir. Med.* 2011; **105**: 106–13.
14 Furuhashi K, Suda T, Nakamura Y, Inui N, Hashimoto D, Miwa S, Hayakawa H, Kusagaya H, Nakano Y, Nakamura H *et al.* Increased expression of YKL-40, a chitinase-like protein, in serum and lung of patients with idiopathic pulmonary fibrosis. *Respir. Med.* 2010; **104**: 1204–10.
15 Berres ML, Papen S, Pauels K, Schmitz P, Zaldivar MM, Hellerbrand C, Mueller T, Berg T, Weiskirchen R, Trautwein C *et al.* A functional variation in CHI3L1 is associated with severity of liver fibrosis and YKL-40 serum levels in chronic hepatitis C infection. *J. Hepatol.* 2009; **50**: 370–6.
16 Ober C, Tan Z, Sun Y, Possick JD, Pan L, Nicolae R, Radford S, Parry RR, Heinzmann A, Deichmann KA *et al.* Effect of variation in CHI3L1 on serum YKL-40 level, risk of asthma, and lung function. *N. Engl. J. Med.* 2008; **358**: 1682–91.
17 Nielsen KR, Steffensen R, Boegsted M, Baech J, Lundbye-Christensen S, Hetland ML, Krintel SB, Johnsen HE, Nyegaard M, Johansen JS. Promoter polymorphisms in the chitinase 3-like 1 gene influence the serum concentration of YKL-40 in Danish patients with rheumatoid arthritis and in healthy subjects. *Arthritis Res. Ther.* 2011; **13**: R109.
18 Bonella F, Cai MT, Ohshimo S, Okuda-Yamamoto A, Guzman J, Costabel U. Serum levels of YKL-40, an interstitial lung disease biomarker in patients with autoimmune alveolar proteinosis (abstract). *Am. J. Respir. Crit. Care Med.* 2010; **181**: A6015.
19 Costabel U, Guzman J, Bonella F, Oshimo S. Bronchoalveolar lavage in other interstitial lung diseases. *Semin. Respir. Crit. Care Med.* 2007; **28**: 514–24.
20 Bonella F, Bauer PC, Griese M, Wessendorf TE, Guzman J, Costabel U. Wash-out kinetics and efficacy of a modified lavage technique for alveolar proteinosis. *Eur. Respir. J.* 2012; **40**: 1468–74.
21 Bonella F, Ohshimo S, Miaotian C, Griese M, Guzman J, Costabel U. Serum KL-6 is a predictor of outcome in pulmonary alveolar proteinosis. *Orphanet J. Rare Dis.* 2013; **8**: 53.
22 Klech H, Pohl W. Technical recommendations and guidelines for bronchoalveolar lavage (BAL). Report of the European Society of Pneumology Task Group. *Eur. Respir. J.* 1989; **2**: 561–85.
23 Kitamura T, Uchida K, Tanaka N, Tsuchiya T, Watanabe J, Yamada Y, Hanaoka K, Seymour JF, Schoch OD, Doyle I *et al.* Serological diagnosis of idiopathic pulmonary alveolar proteinosis. *Am. J. Respir. Crit. Care Med.* 2000; **162**: 658–62.
24 Latzin P, Tredano M, Wust Y, de-Blic J, Nicolai T, Bewig B, Stanzel F, Köhler D, Bahuau M, Griese M. Anti-GM-CSF antibodies in paediatric pulmonary alveolar proteinosis. *Thorax* 2005; **60**: 39–44.
25 Bonella F, Bauer PC, Griese M, Ohshimo S, Guzman J, Costabel U. Pulmonary alveolar proteinosis: new insights from a single-center cohort of 70 patients. *Respir. Med.* 2011; **105**: 1908–16.
26 Uchida K, Nakata K, Suzuki T, Luisetti M, Watanabe M, Koch DE, Stevens CA, Beck DC, Denson LA, Carey BC *et al.* Granulocyte/macrophage colony stimulating factor autoantibodies and myeloid cell immune functions in healthy subjects. *Blood* 2009; **113**: 2547–56.
27 Zheng JL, Lu L, Hu J, Zhang RY, Zhang Q, Chen QJ, Shen WF. Genetic polymorphisms in chitinase 3-like 1 (CHI3L1) are associated with circulating YKL-40 levels, but not with angiographic coronary artery disease in a Chinese population. *Cytokine* 2011; **54**: 51–5.
28 Standardized lung function testing. Official statement of the European Respiratory Society. *Eur. Respir. J. Suppl.* 1993; **16**: 1–100.
29 Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 2005; **21**: 263–5.
30 Li Y, Tian X, Gui Y, Ma A, Li X, Zeng N, Zhang P, Li G, Xu K. Serum markers in patients with idiopathic pulmonary alveolar proteinosis. *Zhonghua Jie He He Hu Xi Za Zhi* 2014; **37**: 497–501.
31 Liu CY, Li Q, Zhou XD, Kolosov VP, Perelman JM. The chitinase-like protein YKL-40 increases mucin5AC production in human bronchial epithelial cells. *Exp. Cell Res.* 2013; **319**: 2866–73.
32 Riabov V, Gudima A, Wang N, Mickley A, Orekhov A, Kzhyshkowska J. Role of tumor associated macrophages in tumor angiogenesis and lymphangiogenesis. *Front. Physiol.* 2014; **5**: 75.
33 Guerra S, Halonen M, Sherrill DL, Venker C, Spangenberg A, Carsin AE, Tarès L, Lavi I, Barreiro E, Martínez-Moratalla J *et al.* The relation of circulating YKL-40 to levels and decline of lung function in adult life. *Respir. Med.* 2013; **107**: 1923–30.
34 Li XY, Donaldson K, Rahman I, MacNee W. An investigation of the role of glutathione in increased epithelial permeability induced by cigarette smoke in vivo and in vitro. *Am. J. Respir. Crit. Care Med.* 1994; **149**: 1518–25.
35 Matsuura H, Hartl D, Kang MJ, Cruz CSD, Koller B, Chupp GL, Homer RJ, Zhou Y, Cho WK, Elias JA *et al.* Role of breast regression protein-39 in the pathogenesis of cigarette smoke-induced inflammation and emphysema. *Am. J. Respir. Cell Mol. Biol.* 2011; **44**: 777–86.
36 Ekaney ML, Otto GP, Sossdorf M, Sponholz C, Boehringer M, Loesche W, Rittirsch D, Wilharm A, Kurzai O, Bauer M *et al.* Impact of plasma histones in human sepsis and their contribution to cellular injury and inflammation. *Crit. Care* 2014; **18**: 543.
37 Campo I, Mariani F, Rodi G, Paracchini E, Tsana E, Piloni D, Nobili I, Kadija Z, Corsico A, Cerveri I *et al.* Assessment and management of pulmonary alveolar proteinosis in a reference center. *Orphanet J. Rare Dis.* 2013; **8**: 40.
38 Punjabi NM, Shade D, Patel AM, Wise RA. Measurement variability in single-breath diffusing capacity of the lung. *Chest* 2003; **123**: 1082–9.
39 Beccaria M, Luisetti M, Rodi G, Corsico A, Zoia MC, Colato S, Pochetti P, Braschi A, Pozzi E, Cerveri I. Long-term durable benefit after whole lung lavage in pulmonary alveolar proteinosis. *Eur. Respir. J.* 2004; **23**: 526–31.
40 Arai T, Inoue Y, Sugimoto C, Inoue Y, Nakao K, Takeuchi N, Matsumuro A, Hirose M, Nakata K, Hayashi S. CYFRA 21-1 as a disease severity marker for autoimmune pulmonary alveolar proteinosis. *Respirology* 2014; **19**: 246–52.
41 Ohshimo S, Ishikawa N, Horimasu Y, Fujitaka K, Haruta Y, Murai H, Hirohashi N, Hattori N, Tanigawa K, Kohno N *et al.* MUC1 gene polymorphisms are associated with serum KL-6 levels and pulmonary dysfunction in pulmonary alveolar proteinosis. *Eur. Respir. J.* 2013; **42**: P2351.

Supplementary Information

Additional supplementary information can be accessed via the *html* version of this article at the publisher's website.

Figure S1 Comparison of baseline serum YKL-40 levels in 34 PAP patients (A) and 50 healthy controls (B) according to smoking habits. Dots represent single patients. Bars represent mean values.
Figure S2 Comparison of baseline serum YKL-40 levels in 34 PAP patients according to disease outcome.
Figure S3 Correlation of baseline serum YKL-40 levels with baseline DL_{CO} (a), PaO_2 (b), AaO_2 (c). Correlation between change in serum YKL-40 and DL_{CO} (d) over time. Dots represent single patients (a-c). Dots represent single measurements (d).
Figure S4 Correlations between serum YKL-40 levels and BALF YKL-40 levels (A), serum LDH levels (B) and BALF lymphocyte counts (C).

