

K. Marten  
V. Dicken  
C. Kneitz  
M. Höhmann  
W. Kenn  
D. Hahn  
C. Engelke

## Interstitial lung disease associated with collagen vascular disorders: disease quantification using a computer-aided diagnosis tool

Received: 26 March 2008  
Revised: 8 June 2008  
Accepted: 24 July 2008  
Published online: 26 August 2008  
© The Author(s) 2008.  
This article is published with open access at Springerlink.com

C. Kneitz  
Dept. of Rheumatology and Clinical Immunology,  
Medizinische Klinik and Poliklinik,  
University Hospital of Würzburg,  
Klinikstr. 6,  
97070 Würzburg, Germany

K. Marten (✉) · C. Engelke  
Department of Radiology,  
University Hospital of Göttingen,  
Robert-Koch-Str. 40,  
37075 Göttingen, Germany  
e-mail: kmarten@med.uni-goettingen.de  
e-mail: c.engelke@med.uni-goettingen.de  
Tel.: +49-551-398965  
Fax: +49-551-399606

M. Höhmann · W. Kenn · D. Hahn  
Department of Radiology,  
University Hospital of Würzburg,  
Josef-Schneider-Str. 2,  
97080 Würzburg, Germany

V. Dicken  
MeVis Research GmbH,  
Universitätsallee 29,  
28359 Bremen, Germany

**Abstract** The purpose of this study was to evaluate a computer-aided diagnosis (CAD) tool compared to human observers in quantification of interstitial lung disease (ILD) in patients with collagen-vascular disorders. A total of 52 patients with rheumatoid arthritis ( $n=24$ ), scleroderma ( $n=14$ ) and systemic lupus erythematosus ( $n=14$ ) underwent thin-section CT. Two independent observers assessed the extent of ILD (EoILD), reticulation (EoRet) and ground-glass opacity (EoGGO). CAD assessed EoILD twice. Pulmonary function tests were obtained. Statistical evaluation used 95% limits of agreement and linear regression analysis. CAD correlated well with dif-

fusing capacity ( $DL_{CO}$ ) ( $R=-0.531$ ,  $P<0.0001$ ) and moderately with forced vital capacity (FVC) ( $R=-0.483$ ,  $P=0.0008$ ). There was close correlation between CAD and the readers (EoILD vs. CAD:  $R=0.716$ ,  $P<0.0001$ ; EoRet vs. CAD:  $R=0.69$ ,  $P<0.0001$ ). Subgroup analysis including patients with minimal EoGGO ( $<15\%$ ) strengthened the correlations between CAD and the readers, readers and PFT, and CAD and PFT. EoILD by readers correlated strongly with  $DL_{CO}$  ( $R=-0.705$ ,  $P<0.0001$ ) and moderately with FVC ( $R=-0.559$ ,  $P=0.0002$ ). EoRet correlated closely with  $DL_{CO}$  and moderately with FVC ( $DL_{CO}$ :  $R=-0.663$ ; FVC:  $R=-0.436$ ;  $P\leq 0.005$ ). The CAD system is a promising tool for ILD quantification, showing close correlation with human observers and physiologic impairment.

**Keywords** Computer-aided diagnosis · Interstitial lung disease · Collagen vascular disorders

### Introduction

In patients with collagen-vascular disorders (CVD) and chronic interstitial lung disease (ILD), formal scoring of disease extent on high-resolution CT (HRCT) has been shown to improve the accuracy of staging, thereby allowing selection of high-risk patients who may benefit from treatment [1]. In patients with scleroderma, for example, increasingly extensive disease on HRCT proved to be a

strong predictor of mortality [2]. In addition, if combined with pulmonary function test data, prognostic information could be obtained in these patients [2]. Furthermore, formal scoring of disease extent plays an important role in therapeutic studies in CVD and may also assist in interpretation of patterns of pulmonary function impairment [1]. This illustrates the need for noninvasive and reproducible scoring systems applicable both to routine clinical practice and the enrolment of patients in pharmaceutical studies.

Recently, computer-aided diagnosis (CAD) has been recognized as a valuable means for improved performance and decision-making due to enhanced detection and evaluation of complex imaging features in the chest [3]. The majority of quantitative analysis methods involve the application of thresholding methods to the segmented pulmonary parenchyma in order to extract regions with attenuation values either above or below a user-defined threshold criterion [3]. This approach may potentially allow a precise, time-efficient, and reproducible quantification of the diseased pulmonary parenchyma, as it requires only minimal user interaction for definition of a threshold value [3, 4]. Hence, the aim of this study was to evaluate the performance of a CAD prototype software for quantification of disease extent in patients with ILD associated with CVD and its correlation with physiological impairment, in comparison with reader-based disease quantification.

## Materials and methods

### Patient recruitment and pulmonary function analyses

Data of a retrospective consecutive patient cohort with known CVD and associated chronic ILD referred for thin-section CT were investigated. The study was approved by the institutional ethics committee. Pulmonary function tests (PFT) were used to measure forced vital capacity (FVC), forced expiratory volume in 1 s ( $FEV_1$ ), and diffusing capacity (carbon monoxide transfer factor,  $DL_{CO}$ ) (Master Screen Body, Viasys Healthcare, Höchberg/ Germany). PFTs were performed by using commonly accepted measurement techniques [8]. Results were expressed as a percentage of predicted performance derived from published standards [5].

### CT acquisition and image interpretation

All CT examinations were acquired with a multidetector-row CT system (Somatom Sensation 16, Siemens Medical Solutions, Erlangen/Germany; 120 kV, 55 effective mAs, 0.75-mm slice collimation, 1-mm reconstruction slice thickness, 1-mm reconstruction increment) at full inspiration in supine position. CT images were reconstructed by using a high spatial frequency algorithm (AB 70) and displayed using lung parenchyma window settings (window level, 500 HU; window width, 1,500 HU).

All the CT images for each patient were independently reviewed by two experienced observers who were blinded to the clinical diagnosis and PFT results. The observers were asked to score each lobe for the extent of normal lung and interstitial lung disease (adding up to 100%) to the nearest 5% counting the lingula as a separate lobe. Interstitial lung disease was subclassified

into reticulation and ground-glass opacification (GGO), adding up to 100%. If a reticular pattern was reported, the coarseness of the reticulation was graded on a 3-point scale (grade 1=intralobular, grade 2=microcystic with cysts 1–3 mm in size, grade 3=macrocytic with cysts >3 mm in size). Values for extent of ILD and morphological parameters were averaged across all lobes.

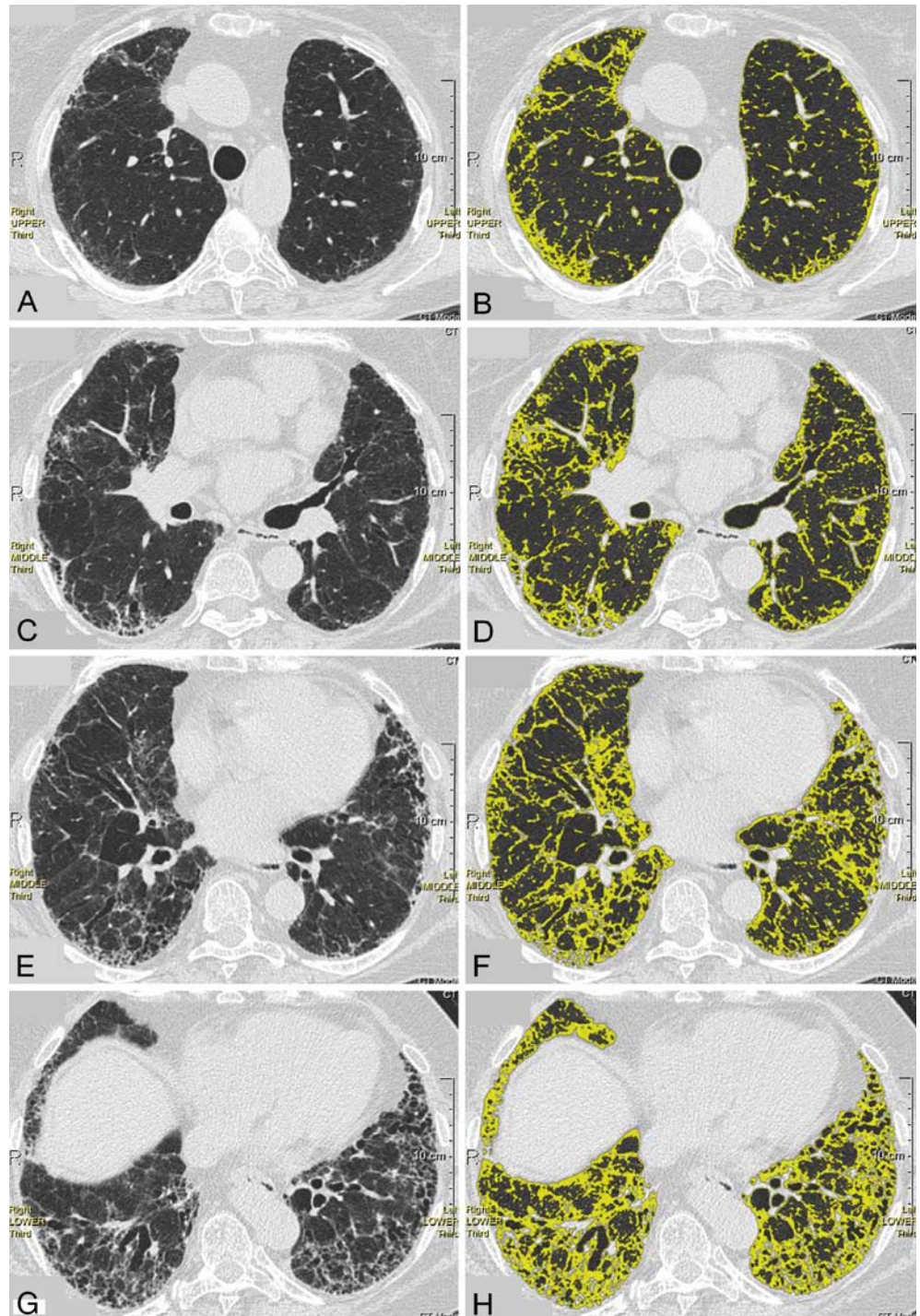
### CT evaluation by the CAD system

The CAD software (MeVisPULMO 3D 3.11, MeVIS Research, Bremen, Germany) is a prototype system designed for CT-based automated segmentation of the airways, lungs, and pulmonary lobes. The algorithm has been described previously [4] and is characterized by an initial segmentation of the airways, based on a region-growing method, that is followed by airway analysis in order to obtain a graphic representation of the bronchial tree, including information about direction, diameter, and length of each branch [6, 7]. The bronchi belonging to each pulmonary segment are identified automatically.

Based on the bronchial tree, a preliminary mask of the lung parenchyma is created using a conventional region-growing method with a fixed threshold interval (–1,024 to –400 HU) [4]. Subsequently, a marker-based 3D watershed transform is used that is confined to the area of the preliminary lung masks in order to obtain separate masks for each lung [4, 8]. The segmentation of the pulmonary vasculature is restricted to the area defined by the previously computed lung masks and is performed using a conventional 3D region-growing algorithm, thereby creating a superset of all larger pulmonary vessels. Accordingly, lobar segmentation is based on the absence of larger vessels in proximity to the lobar boundaries [8]. For this purpose, a Euclidean distance transform is used, and lobar areas are obtained from the preprocessed image using a multidimensional interactive watershed transform [9]. These region identification steps enable extraction of various CT parameters [4]. Initially developed for quantification of pulmonary emphysema, the software allows extraction of the volume, mean density, pixel index, bulla index, and emphysema type on a lobar or segmental basis [4]. Likewise, it is feasible to conduct threshold-based segmentation of the percentage of areas with interstitial lung disease, which in this context is defined as an area of high attenuation. For this purpose, the threshold is defined by the user. The current prototype software does not correct for gravity-dependent opacity or overperfusion.

For the current study, all CT images for each patient were transferred from the CT system to a dedicated storage workstation (Fujitsu Siemens Celsius, Siemens Medical Solutions, Erlangen/Germany). Each set of CT images for each patient was evaluated twice using the MeVisPULMO 3D software by application of a threshold value of –800 HU (Figs. 1, 2 and 3). This threshold value was chosen

**Fig. 1** A 74-year-old female patient with rheumatoid arthritis. Thin-section CT (**a, c, e, g**) shows extensive interstitial lung disease with predominantly subpleural and basal reticulation, ground-glass opacification and associated traction bronchiectasis indicating fibrosis. These areas are precisely segmented using CAD (**b, d, f, h**; high-attenuation areas marked in yellow)

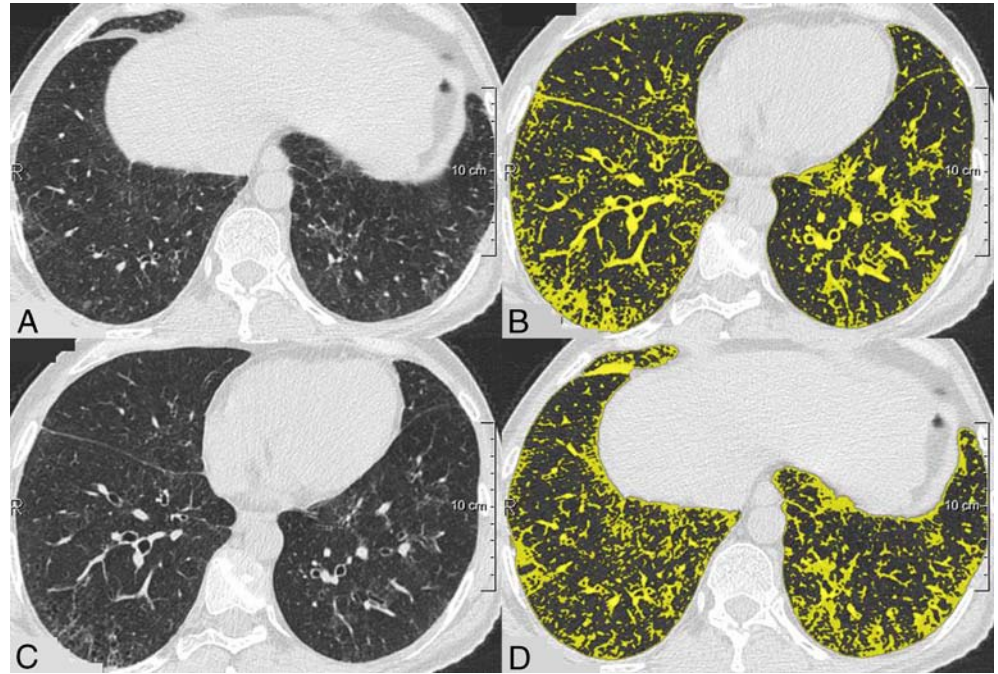


after separate data from a previous patient collective with ILD had demonstrated an optimum correlation between the extent of high-attenuation areas and diffusing capacity (data not shown). As interstitial lung disease tends to be widespread and, in contrast to emphysema, does not tend

to be confined to single lobes or segments, total extent of ILD was reported (expressed as percentage areas with high attenuation, i.e., attenuation above the defined threshold value).



**Fig. 2** A 53-year-old male patient with rheumatoid arthritis. Thin-section CT (**a, c**) displays areas of ground-glass opacification and associated subpleural reticulation with a basal predominance. These findings are depicted as high-attenuation areas by the CAD application (**b, d**)

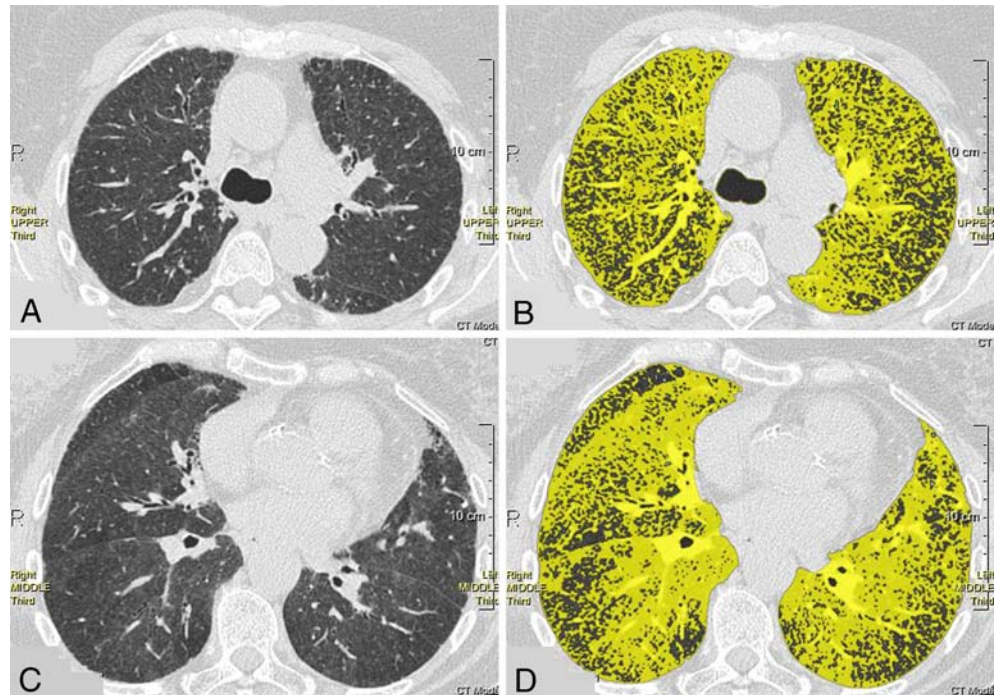


### Statistical evaluation

Data evaluation was performed using the spreadsheet-based statistical software package StatsDirect (Camcode, Herts, England, version 2.6.5). Measurement reproducibility of repeated CAD assessments as well as interobserver agreement between the two study readers was tested using

the method of Bland and Altman, which calculates 95% limits of agreement between two measurements [10]. This value is an expression of 95% of all measurements that can be expected to lie within the calculated (limits of agreement) range. Mean values of both readers were used for regression analyses. The relationships among the CAD, the readers and the PFT results were calculated using linear

**Fig. 3** A 70-year-old female patient with systemic lupus erythematosus and known chronic interstitial lung disease. Thin-section CT (**a, c**) displays widespread subtle ground-glass opacification with associated traction bronchiolectasis indicating fine fibrosis, as well as a discrete subpleural reticular pattern. CAD enhances visualization of the disease extent (**b, d**)



regression analysis and Pearson's product moment correlation. Linear regression parameters were calculated by the least squares method. Pearson's product moment correlation coefficient was given as a measure of linear association between two variables:  $r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}$ .

Calculations were first performed for all 52 patients and then for a reduced subset of patients with GGO scores below 15% to evaluate the effect of the amount of GGO on the correlations among CAD, readers, and PFTs. Further, a stratum for the three patient diagnoses was introduced using dummy variables (representing a categorical variable with more than two levels in the regression model) and subsequent multiple linear regression analysis to test the effect of the patient diagnosis on the relationship between the disease extent as measured by CAD and readers, and the pulmonary functional impairment as measured by PFTs [11]. A *P* value of less than 0.05 was accepted to indicate statistical significance.

## Results

The study group comprised 52 consecutive patients (14 male, 36 female, mean age  $59 \pm 13$  years) with known CVD and chronic pulmonary disease. Patients had received a diagnosis of rheumatoid arthritis (RA,  $n=24$ ), scleroderma (PSS,  $n=14$ ), or systemic lupus erythematosus (SLE,  $n=14$ ) that met the respective diagnostic criteria of the American College of Rheumatology [12]. The average duration of CVD was  $10.27 \pm 7.04$  years in patients with PSS,  $11.25 \pm 9.60$  years in patients with RA, and  $11.0 \pm 6.65$  years in patients with SLE. The average duration of respiratory symptoms was  $9.09 \pm 7.46$  years in patients with PSS,  $3.79 \pm 2.57$  years in patients with RA, and  $6.9 \pm 3.84$  years in patients with SLE. None of the patients reported a history of smoking or smoked at the time of presentation. Lung biopsies were not performed in any of the patients. All patients displayed crackles on physical examination, complained of respiratory symptoms (dyspnea, cough), and underwent treatment with a steroid or cytotoxic agent according to various treatment regimens at the time of CT. At the time of presentation, no patient displayed clinical or laboratory signs of infection. None of the patients had overt clinical or echocardiographic evidence of pulmonary arterial hypertension. The mean time interval between PFTs and thin-section CT was  $1.5 \pm 2.5$  days (range: 0–7 days). On PFT, average  $DL_{CO}$  was  $58.2 \pm 15.1\%$  of predicted, average FVC was  $86.1 \pm 23.5\%$  of predicted, and average  $FEV_1$  was  $87.2 \pm 21.0\%$  of predicted.

## Measurement reproducibility and interobserver agreement

There was total concordance between the first and second measurements of the high-attenuation areas (HAV) by the CAD tool corresponding to the extent of ILD (95% limits of agreement=0 to 0, intra-class correlation coefficient=1). The interobserver agreement of both readers on the extent of ILD and the various morphological patterns was good (95% limits of agreement=–27.00 to 17.03%, –24.51 to 15.48%, –32.35 to 20.43%, and –1.10 to 0.82 for extent of ILD, extent of reticulation, extent of GGO, and coarseness of reticulation, respectively; intraclass correlation coefficient=0.89, 0.87, 0.70, and 0.61 for extent of ILD, extent of reticulation, extent of GGO, and coarseness of reticulation, respectively).

## Correlation between CAD results and PFTs

The percentage of high-attenuation areas (average HAV=  $25.0 \pm 16.9\%$ ), corresponding to the extent of ILD by CAD, showed a significant correlation with  $DL_{CO}$  ( $R=-0.531$ ; 95% CI=–0.706 to –0.293;  $P<0.0001$ ) and FVC ( $R=-0.483$ ; 95% CI=–0.680 to –0.221;  $P=0.0008$ ), but no significant correlation with  $FEV_1$  (Table 1).

## Correlation between reader results and PFTs

On thin-section CT, all patients displayed findings of ILD (ILD detected by the readers: average extent of ILD=  $36.3 \pm 27.2\%$ , average extent of reticulation=  $27.0 \pm 23.3\%$ , average extent of GGO=  $9.2 \pm 17.0\%$ , average coarseness of a reticular pattern=  $1.1 \pm 0.6\%$ ). Major respiratory or cardiac motion artefacts were not noted in any of the patients. Average extent of ILD was correlated closely with  $DL_{CO}$  ( $R=-0.705$ ; 95% CI=–0.831 to –0.511;  $P<0.0001$ ). There was a moderate correlation of the average extent of ILD and FVC ( $R=-0.559$ ; 95% CI=–0.742 to –0.299;  $P=0.0002$ ), and a weak but significant correlation between the average extent of ILD and  $FEV_1$  ( $R=-0.379$ ; 95% CI=–0.615 to –0.08;  $P=0.014$ ). The extent of reticulation correlated closely with  $DL_{CO}$  and moderately with FVC ( $DL_{CO}$ :  $R=-0.663$ ; 95% CI=–0.805 to –0.449;  $P<0.0001$ ; FVC:  $R=-0.436$ ; 95% CI=–0.658 to –0.144;  $P=0.005$ ), but there was no significant correlation between the extent of reticulation and  $FEV_1$  (see Table 1). The extent of GGO did not show any significant correlation with any of the PFTs (Table 1). The average coarseness correlated moderately with  $DL_{CO}$  ( $R=-0.435$ ; 95% CI=–0.653 to –0.151;  $P=0.004$ ), but not with FVC or  $FEV_1$  (Table 1).

**Table 1** Correlation between reader/CAD results and PFTs

		DL <sub>CO</sub>			FVC			FEV <sub>1</sub>		
		R	95% CI	P	R	95% CI	P	R	95% CI	P
CAD	Extent of HAV	-0.531	-0.706 to -0.293	<0.0001*	-0.483	-0.680 to -0.221	0.0008*	-0.249	-0.498 to 0.038	0.088
Readers	Extent of ILD	-0.705	-0.831 to -0.511	<0.0001*	-0.559	-0.742 to -0.299	0.0002*	-0.379	-0.615 to -0.08	0.014*
	Extent of reticulation	-0.663	-0.805 to -0.449	<0.0001*	-0.436	-0.658 to -0.144	0.005*	-0.241	-0.511 to 0.071	0.128
	Extent of ground-glass opacification	-0.234	-0.502 to 0.075	0.136	-0.304	-0.563 to 0.008	0.056	-0.276	-0.538 to 0.034	0.081
	Coarseness	-0.435	-0.653 to -0.151	0.004*	-0.191	-0.474 to -0.128	0.0238	-0.036	-0.340 to 0.274	0.822

Statistically significant correlations marked with an *asterisk*

R Correlation coefficient, 95% CI confidence interval (Fisher's z transformed)

### Correlation between CAD and reader results

There was a close correlation between the extent of ILD by the readers and CAD ( $R=0.716$ ; 95% CI=0.529–0.836;  $P<0.0001$ ). The extent of reticulation as assessed by the readers correlated closely with CAD ( $R=0.69$ ; 95% CI=0.492–0.821;  $P<0.0001$ ). The average coarseness as determined by the readers correlated moderately with CAD ( $R=0.508$ ; 95% CI=0.245–0.701;  $P=0.0005$ ). However, there was no significant correlation between the extent of GGO as assessed by the readers and the extent of ILD by CAD (Table 2).

### Correlation between CAD results and PFTs in a subgroup of patients with minimal ground-glass opacification

In the 34 patients with minimal GGO, the correlations between CAD and PFTs were further improved (Table 3). There were significant correlations between CAD and DL<sub>CO</sub> ( $R=-0.56$ ; 95% CI=-0.758 to -0.269;  $P=0.0007$ )

as well as between CAD and FVC ( $R=-0.521$ ; 95% CI=-0.728 to -0.228;  $P=0.001$ ), whereas the correlation between CAD and FEV<sub>1</sub> remained insignificant (Table 3).

### Correlation between reader results and PFTs in a subgroup of patients with minimal ground-glass opacification

When the analysis was restricted to a patient subgroup of 34 patients with minimal (<15%) GGO, the correlations between reader results and DL<sub>CO</sub> were improved (Table 3). In detail, there were close correlations between the extent of ILD and DL<sub>CO</sub> ( $R=-0.702$ ; 95% CI=-0.842 to -0.472;  $P<0.0001$ ) and between the extent of reticulation and DL<sub>CO</sub> ( $R=-0.69$ ; 95% CI=-0.836 to -0.455;  $P<0.0001$ ), and a moderate correlation between the average coarseness and DL<sub>CO</sub> ( $R=-0.543$ ; 95% CI=-0.747 to -0.246;  $P=0.001$ ). In addition, the correlation between the extent of reticulation and FVC was improved ( $R=-0.462$ ; 95% CI=-0.705 to -0.121;  $P=0.01$ ).

**Table 2** Correlation between reader and CAD results

		CAD: HAV		
		R	95% CI	P
Readers	Extent of ILD	0.716	0.529 to 0.836	<0.0001*
	Extent of reticulation	0.690	0.492 to 0.821	<0.0001*
	Extent of ground-glass opacification	0.199	-0.107 to 0.472	0.199
	Coarseness	0.508	0.245 to 0.701	0.0005*

Statistically significant correlations marked with an *asterisk*

R Correlation coefficient, 95% CI confidence interval (Fisher's z transformed)



**Table 3** Correlation between reader/CAD results and PFTs in subgroup of patients with minimal ground-glass opacification

		DL <sub>CO</sub>			FVC			FEV <sub>1</sub>		
		R	95% CI	P	R	95% CI	P	R	95% CI	P
CAD	Extent of HAV	-0.56	-0.758 to -0.269	0.0007*	-0.521	-0.728 to 0.228	0.001*	-0.316	-0.574 to -0.001	0.051
Readers	Extent of ILD	-0.702	-0.842 to -0.473	<0.0001*	-0.488	-0.721 to -0.155	0.006*	-0.288	-0.583 to 0.073	0.116
	Extent of reticulation	-0.690	-0.836 to -0.455	<0.0001*	-0.462	-0.705 to -0.121	0.010*	-0.255	-0.559 to 0.109	0.166
	Coarseness	-0.543	-0.747 to -0.246	0.001*	-0.248	-0.558 to 0.123	0.186	-0.062	-0.408 to 0.299	0.74

Statistically significant correlations marked with an *asterisk*

R Correlation coefficient, 95% CI confidence interval (Fisher's z transformed)

#### Correlation between CAD and reader results in a subgroup of patients with minimal ground-glass opacification

Restricting the analysis to the subgroup of 34 patients with minimal GGO also improved the correlations between reader and CAD results (Table 4). There were close correlations between the extent of ILD by the readers and CAD ( $R=0.779$ ; 95% CI=0.596–0.886;  $P<0.0001$ ), as well as between the extent of reticulation and CAD ( $R=0.722$ ; 95% CI=0.504–0.854;  $P<0.0001$ ), and a moderate correlation between the average coarseness and CAD ( $R=0.594$ ; 95% CI=0.314–0.778;  $P=0.0003$ ).

#### Effect of patient diagnosis on correlation between extent of ILD and DL<sub>CO</sub>

On multivariate linear regression analysis there was no significant influence of any of the patient diagnoses (RA, SLE and PSS) on the correlation between the extent of ILD by CAD and the DL<sub>CO</sub> [for correlation of extent of ILD by CAD/disease entity with DL<sub>CO</sub>:  $F=18.42$ ; for CAD:  $R=-0.508$  (95% CI=-0.706 to -0.293,  $P=0.0003$ ); for the diagnoses RA, SLE, or PSS:  $P=0.16-0.22$ ]. Furthermore, there was no significant influence of any of the patient diagnoses on the correlation between the extent of ILD by the readers and the DL<sub>CO</sub> [for correlation of extent of ILD by

readers/disease entity with DL<sub>CO</sub>:  $F=39.62$ ; for readers:  $R=-0.701$  (95% CI=-0.831 to -0.511,  $P<0.0001$ ); for diagnoses RA, SLE, or PSS:  $P=0.08-0.25$ ]. Moreover, the strength of the univariate correlation between the extent of ILD by CAD (Table 1) or by the readers and the DL<sub>CO</sub> was not reduced substantially by the disease stratification ( $P>0.05$ ).

#### Discussion

Formal scoring of disease extent on HRCT in patients with ILD associated with CVD has been performed by several investigators [13–16]. The extent of disease on HRCT has been found to be inversely correlated to arterial oxygen levels at rest, and extensive pulmonary fibrosis correlates with the presence of neutrophilia on bronchoalveolar lavage [14, 17]. In patients with scleroderma, there is a close correlation between the extent of disease on HRCT and diffusing capacity [14]. More recently, a combined staging system including both disease extent on HRCT (based on the evaluation of five predefined anatomic sections scored to the nearest 5%) as well as pulmonary function test data has been shown to provide discriminatory prognostic information in these patients [2]. In addition, formal scoring of disease extent plays an important role in therapeutic studies in CVD. In a recent study in patients with scleroderma, the evaluation of disease extent at

**Table 4** Correlation between reader/CAD results in subgroup of patients with minimal ground-glass opacification

		CAD: HAV		
		R	95% CI	P
Readers	Extent of ILD	0.779	0.596–0.886	<0.0001*
	Extent of reticulation	0.722	0.504–0.854	<0.0001*
	Coarseness	0.594	0.314–0.778	0.0003*

Statistically significant correlations marked with an *asterisk*

R Correlation coefficient, 95% CI = confidence interval (Fisher's z transformed)

baseline strengthened the treatment effect of oral cyclophosphamide against placebo [1, 18]. Furthermore, the evaluation of disease extent may assist in interpretation of patterns of pulmonary function impairment. In rheumatoid arthritis, for example, knowledge of the extent of pulmonary fibrosis may aid in ascribing reduced lung volumes to interstitial or pleural disease [1].

However, as there are currently no standardized scoring systems for the assessment of disease extent on thin-section CT of patients with ILD associated with CVD, the application of CAD systems offers considerable promise. Most quantitative analytic methods involve the application of thresholding methods to the segmented pulmonary parenchyma in order to extract regions with attenuation values either above or below a user-defined threshold criterion [3]. Alternatively, they focus on the generation of histograms reflecting a regional or global distribution of pixel attenuations [3]. To date, few studies have investigated the quantitative CAD assessment of ILD in patients with pulmonary fibrosis using computer-derived histogram indices, showing correlation between different histogram features and physiologic impairment [19–22]. In addition, attempts have been made to better quantify complex lung patterns by application of texture analysis using measurements such as entropy, or by using the “Adaptive Multiple Feature Method” which identifies normal lung as well as different patterns of infiltrative lung disease [3, 23].

The prototypical software application MeVisPULMO used in our study was designed to assist the radiologist in functional analysis of thoracic CT data. It allows extraction of volumes and CT parameters such as mean attenuation, percentage of high or low attenuation areas, or pixel index on a lobar level [4]. Preliminary studies using the CAD system in patients undergoing lung volume reduction surgery showed that the CT-based prediction of FEV<sub>1</sub> correlated significantly with perfusion scintigraphy results [4]. However, the MeVisPULMO software was not evaluated in quantification of ILD.

In our study, the percentage of segmented lung by CAD correlated significantly with the extent of ILD as evaluated by the readers and further with morphologic parameters including the extent of reticulation and the degree of coarseness (Figs. 1, 2). Although the correlation of extent of ILD by CAD with diffusing capacity was slightly weaker than that between diffusing capacity and the extent of ILD by the readers, the CAD-functional correlation is comparable with the results of human observers in other studies [16, 24]. In concordance with the readers, the CAD result closely correlated with DL<sub>CO</sub> and the FVC. Diffusing capacity reflected the extent of ILD on CT more accurately than the other PFTs, which is in keeping with the results of Wells et al. in patients with systemic sclerosis [14]. Notably, the results of the multivariate analyses in our study showed that these relationships were independent of the disease entity in the individual cases, indicating a robust performance of the CAD tool in this heterogeneous patient cohort.

The scoring system used in our study was based on lobar assessment of total disease extent, reticulation, GGO and coarseness (modified after ref. [25]). We observed good correlations between the extent of ILD as well as the extent of reticulation and the DL<sub>CO</sub> or FVC. The lack of correlation between the extent of GGO and DL<sub>CO</sub> is likely to be partly due to the limited extent of this pattern in the study group. However, the correlations between CAD and readers (extent of disease, extent of reticulation, coarseness) and PFTs (DL<sub>CO</sub>, FVC) were strengthened when the analyses were restricted to patients with minimal GGO (<15%). In collagen-vascular disorders, GGO may reflect inhomogeneous lung ventilation or perfusion as well as infiltrative (e.g., infectious) disease or drug toxicity [16, 26, 27]. GGO mixed with a reticular pattern is indicative of fine fibrosis below the spatial resolution of thin-section CT, particularly if traction bronchiolectasis is present within these areas (Fig. 3) [26]. The diversity of causes of GGO may have contributed to the weak relationships between CAD and PFTs, CAD and the readers, as well as between the readers and PFTs in patients with extensive GGO.

A number of potential limitations of our study need to be considered. Firstly, our study focused on quantification of total disease extent, and no pattern-based analysis of the CT examinations was performed. However, the disease extent was concordantly shown in this and in previous series to be a strong predictor of functional pulmonary impairment and therefore the stage of pulmonary disease, which is an important prerequisite for evaluation of a CAD tool [13, 14]. An issue that needs consideration for optimization of the CAD tool is the overlap of attenuation values of ILD and pulmonary vasculature, leading to inclusion of small pulmonary vessels, which are likely to have resulted in an overestimation of disease extent. Furthermore, the quantification of disease extent using CAD was limited to high-attenuation areas. Thus, its performance is based on a density dichotomy between the normal parenchyma and fibrotic lung areas, disregarding areas of low attenuation potentially reflecting a bronchiolitic component. However, CAD quantification of small-airways disease associated with CVD would strongly depend on the use of expiratory CT examinations, which were not available in our patient cohort. The CAD system used does not allow differentiation of reticular and amorphous elements of increased attenuation, which represents a practical limitation, as does the inclusion of relatively overperfused areas in subjects with coexisting small-airways disease, as well as the lack of correction for gravity-dependent opacities. Finally, CAD evaluation did not focus on anatomic compartments of the lung; however, in contrast to emphysema, ILD tends to be widespread and both morphologically and functionally usually does not reflect a lobar distribution pattern. Formal HRCT scoring, particularly in pharmaceutical trials, is commonly performed using predefined anatomic levels rather than pulmonary lobes, as discontinuous HRCT examinations



(as opposed to volumetric multidetector-row CT examinations) are still widely performed for reasons of radiation protection. The use of CAD in the evaluation of single interspaced thin sections should therefore be subject to further investigation.

In conclusion, the CAD application used proved to be time-efficient, requiring a mean evaluation time of less than 5 min, and to be well reproducible to aid in significantly reducing interobserver variability. It holds promise to become a valuable tool for quantification of interstitial lung

disease, showing close correlation with human observers and physiologic impairment. These observations should be confirmed in future studies enrolling larger cohorts of patients with ILD.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

## References

- Devaraj A, Wells AU, Hansell DM (2007) Computed tomographic imaging in connective tissue diseases. *Semin Respir Crit Care Med* 28:389–397
- Goh NS, Desai SR, Veeraraghavan S, Hansell DM, Copley SJ et al (2008) Interstitial lung disease in systemic sclerosis: a simple staging system. *Am J Respir Crit Care Med* 177:1248–1254
- Ko JP, Naidich DP (2004) Computer-aided diagnosis and the evaluation of lung disease. *J Thorac Imaging* 19:136–155
- Kuhnigk J-M, Dicken V, Zidowitz S et al (2005) New tools for computer assistance in thoracic CT. Part 1: functional analysis of lungs, long lobes, and bronchopulmonary segments. *Radiographics* 25:525–536
- European Respiratory Society (1993) Standardized lung function. *Eur Resp J* 6:49
- Selle D, Preim B, Schenk A, Peitgen HO (2002) Analysis of vasculature for liver surgical planning. *IEEE Trans Med Imaging* 21:1344–1357
- Selle D, Peitgen HO (2001) Analysis of the morphology and structure of vessel systems using skeletonization. In: Chen CT, Clough AV (eds) *Proceedings of SPIE: medical imaging 2001. Physiology and function from multidimensional images*, vol 4321. International Society for Optical Engineering, Bellingham, WA, pp 271–281
- Kuhnigk JM, Hahn HK, Hindennach M et al (2003) Lung lobe segmentation by anatomy-guided 3D watershed transform. In: Sonka M, Fitzpatrick JM (eds) *Proceedings of SPIE: medical imaging 2003. Image processing*, vol 5032. International Society for Optical Engineering 2003, Bellingham, WA, pp 1482–1490
- Hahn HK, Peitgen HO (2003) IWT-interactive watershed transform: a hierarchical method for efficient interactive and automated segmentation of multi-dimensional gray-scale images. In: Sonka M, Fitzpatrick JM (eds) *Proceedings of SPIE: medical imaging 2003. Image processing*, vol 5032. International Society for Optical Engineering, Bellingham, WA, pp 643–653
- Bland M, Altman DG (1986) Statistical methods for assessing the difference between two methods of measurement. *Lancet* i:307–310
- Kleinbaum DG et al (1998) *Applied regression analysis and other multivariable methods*, 3rd edn. Duxbury, Belmont, CA
- Homepage of the American College of Rheumatology (2008) <http://www.rheumatology.org/publications/classification>. Accessed 13 March 2008
- Desai SR, Verraraghavan S, Hansell DM et al (2004) CT features of lung disease in patients with systemic sclerosis: comparison with idiopathic pulmonary fibrosis and non-specific interstitial pneumonia. *Radiology* 232:560–567
- Wells AU, Hansell DM, Rubens MB et al (1997) Fibrosing alveolitis in systemic sclerosis: indices of lung function in relation to extent of disease on computed tomography. *Arthritis Rheum* 40:1229–1236
- Bankier AA, Kiener HP, Wiesmayr MN et al (1995) Discrete lung involvement in systemic lupus erythematosus: CT assessment. *Radiology* 196:835–840
- Biederer J, Schnabel A, Muhle C et al (2004) Correlation between HRCT finding, pulmonary function tests and bronchoalveolar lavage cytology in interstitial lung disease associated with rheumatoid arthritis. *Eur Radiol* 14:272–280
- Wells AU, Hansell DM, Rubens MB et al (1997) Fibrosing alveolitis in systemic sclerosis: bronchoalveolar lavage findings in relation to computed tomographic appearances. *Am J Respir Crit Care Med* 150:462–468
- Tashkin DP, Elashoff R, Clements PJ, et al for the Scleroderma Lung Study Research Group (2006) Cyclophosphamide versus placebo in scleroderma lung disease. *N Engl J Med* 354:2655–2666
- Best AC, Lynch AM, Bozic CM et al (2003) Quantitative CT indexes in idiopathic pulmonary fibrosis: relationship with physiologic impairment. *Radiology* 228:407–414
- Hartley PG, Galvin JR, Hunninghake GW et al (1994) High-resolution CT-derived measures of lung density are valid indexes of interstitial lung disease. *J Appl Physiol* 76:271–277
- Behr J, Mehnert F, Beinert T et al (1992) Evaluation of interstitial lung disease by quantitative high-resolution computed tomography. *Am Rev Respir Dis* 145(suppl):A191
- Rienmüller RK, Behr J, Kalender WA et al (1991) Standardized quantitative high resolution CT in lung disease. *J Comput Assist Tomogr* 15:742–749
- Uppaluri R, Hoffman EA, Sonka M et al (1999) Computer recognition of regional lung disease patterns. *Am J Respir Crit Care Med* 117:991–998
- Terasaki H, Fujimoto K, Hayabuchi N et al (2004) Respiratory symptoms in rheumatoid arthritis: relation between high resolution CT findings and functional impairment. *Rad Med* 3:179–185
- Gay SE, Kazerooni EA, Toews GE et al (1998) Idiopathic pulmonary fibrosis: predicting response to therapy and survival. *Am J Respir Crit Care Med* 157:1063–1072
- Remy-Jardin M, Giraud R, Remy J et al (1993) Importance of ground-glass attenuation in chronic diffuse infiltrative lung disease: pathologic-CT correlation. *Radiology* 189:693–698
- Ellis SM, Hansell DM (2002) Idiopathic interstitial pneumonias: imaging-pathology correlation. *Eur Radiol* 12:610–626