Online Appendices:

CT Protocols

The CT scans were obtained using a 64-slice multiple detector CT scanner (Somatom Sensation 64, Siemens, Erlangen, Germany) or a 4-slice multiple detector CT scanner (Siemens Volume Zoom, Siemens, Erlangen, Germany. To satisfy requirements for processing by the CALIPER algorithm, all scans were reconstructed using a high spatial frequency, B70 kernel (Siemens, Munich, Germany). All patients were scanned from lung apices to bases, supine, at full inspiration, with 1·0mm section thicknesses using a peak voltage of 120kVp with tube current modulation (range 30-140 mA). Images were viewed at window settings optimized for the assessment of the lung parenchyma (width 1500 H.U.; level -500 H.U.).

Pulmonary function tests:

Pulmonary function tests were analyzed if performed within 3 months of the corresponding CT scan according to established protocols [1]. Spirometry (Jaeger Master screen PFT, Carefusion Ltd., Warwick, UK), plethysmographic lung volumes (Jaeger Master screen Body, Carefusion Ltd., Warwick, UK), and diffusion capacity for carbon monoxide (DLco) (Jaeger Master screen PFT, Carefusion Ltd., Warwick. UK) Parameters assessed: forced expiratory volume in one second (FEV1), forced vital capacity (FVC), total lung capacity (TLC), transfer coefficient of the lung for carbon monoxide (Kco) and single breath carbon monoxide diffusing capacity corrected for hemoglobin concentration (DLco). The composite physiologic index (CPI) was

calculated using the formula: 91·0 - (0·65 x % predicted DLco) - (0·53 x % predicted

FVC) + (0·34 x % predicted FEV1). [2]

Echocardiography:

A measure of pulmonary hypertension was obtained from right ventricular systolic

pressure (RVSP) evaluated on transthoracic echocardiography, and only included if

performed within 3 months of the CT scan.

Consensus formulation for visual scores:

The identification of systematic biases in visual scores was achieved by plotting the

spread of differences in parenchymal pattern scores between observers. The most

disparate 5% (two standard deviations) of values were arbitrated by a third scorer

for all parameters except traction bronchiectasis, thereby minimizing bias within the

original scorers. In total, 32% of all IPF cases had at least one parenchymal pattern

score arbitrated by the third scorer. The original scorers derived a consensus for the

traction bronchiectasis score. If a single parenchymal subtype extent was changed at

consensus, the other parameters were modified, following CT review, to retain an

overall sum of 100% for the four parenchymal subtypes. Similarly, if the lobar

percentages of total interstitial disease, emphysema or mosaicism varied, the other

two parameter extents were rescored. The final aorta and pulmonary artery

measurements represented an average of the individual two scorers measurements.

Visual CT Variable (n = 283)	Single determination standard deviation
CT Interstitial lung disease extent	7·24
CT Ground glass opacity	6·15
CT Reticular pattern	5·24
CT Honeycombing	7·88
CT Consolidation	2·69
CT Total emphysema	4·99
CT Mosaic attenuation	3·83
CT Traction bronchiectasis severity	1·43

Supplementary Table 1. Single determination standard deviation values of visual CT scores for idiopathic pulmonary fibrosis cases. CT = computed tomography.

Evaluation of parenchymal vessel volume subdivisions

To interrogate the PVV in greater detail, subdivisions of PVV were analysed. Structures classified as vessels were separated based on their cross-sectional area (CSA) measured on axial CT slices. Vessels measuring <5mm^2, <10mm^2, <15mm^2, <20mm^2, over 20mm^2 CSA (Supplementary Figures 1+2) as well as vessels over 5mm^2 CSA were considered in separate analyses. Vessel volume was derived after correcting for z-axis CT slice thickness and vessel volume was expressed as a percentage of the total lung volume as measured by CALIPER.

Relationships between PVV subdivisions and DLco and PVV were evaluated using linear regression analysis (Supplementary Table 2). Similar linkages with DLco were identified for <5mm^2 PVV CSA (PVV5), <10mm^2 PVV CSA (PVV10), <15mm^2 PVV CSA (PVV15) and <20mm^2 PVV CSA (PVV20). To exclude the possibility that the PVV linkage with DLco could be secondary to contamination from misclassified fine

reticular pattern in the PVV score, structures <5mm^2 in CSA were excluded from the PVV score. The resulting score considered structures between over 5mm^2 as shown in Supplementary Table 2 (PVV>5), and demonstrated comparable linkages for DLco as other vessel subdivisions.

On mortality analyses, the <5mm^2 and <10mm^2 PVV thresholds were shown to have the highest hazard ratios (Supplementary Table 3). The results are in keeping with previous investigations in patients with chronic obstructive pulmonary disease that demonstrated good correlations between a 5mm^2 CSA vessel size threshold and DLco. [3, 4] When structures <5mm^2 were excluded from analysis (thereby excluding the majority of fine reticular pattern), the PVV>5 score remained strongly predictive of mortality, suggesting that the mortality signal related to PVV is not predominantly influenced by misclassification of fibrotic features such as fine reticulation.

In a second validation step, PVV>5 was substituted for PVV in multivariate models. When CALIPER variables alone were evaluated in a multivariate Cox proportional hazards model, only PVV>5 and honeycombing were independently predictive of mortality. A composite variable to predict mortality was derived using the following formula:

CALIPER-PVV>5-HC score = (PVV>5 x 78·4524) + (CALIPER honeycombing x 14·5523)

When PVV>5 was substituted into the model containing CALIPER and visual CT variables and pulmonary function tests, PVV>5 retained independent prognostic

significance and a composite variable to predict mortality was derived using the following formula:

$$\text{CALIPER-PVV>5-HC-CPI score} = (\text{PVV>5} \times 36{\cdot}7190) + (\text{CALIPER honeycombing} \times 20{\cdot}0476) + (\text{CPI} \times 4{\cdot}5441)$$

Both the CALIPER-PVV>5-HC score and the CALIPER-PVV>5-HC-CPI score were converted into categorical scores by aligning the individual scores in ascending numerical order and dividing the respective cohorts into three equally sized groups (n=83). Both models were analysed alongside the GAP score (Table 4) and the GAP index staging system (Table 5), and demonstrated similar results to the models containing PVV. The CALIPER-PVV>5-HC-CPI model was more strongly predictive of mortality than the GAP index staging system, with the results confirmed on bootstrapping of 1000 samples (Table 5). The CALIPER-PVV>5-HC model was equivalent to the GAP index staging system with regard to mortality prediction on bootstrapping of 1000 samples (Table 5).

Supplementary Figure 1. Coronal renderings of pulmonary vessel volume separated according to vessel cross-sectional area on axial computer tomography images in a 75-year-old male patient with idiopathic pulmonary fibrosis. A) Vessels <5mm^2 in cross-sectional area (CSA) (red). B) Vessels <10mm^2 CSA (green). C) Superimposed vessels of <5mm^2 and <10mm^2 CSA. D+E) Five vessel size subdivisions superimposed on coronal (D) and sagittal (E) renderings. Vessels <15mm^2 CSA (yellow); Vessels <20mm^2 CSA (blue); Vessels >20mm^2 CSA (purple).

Supplementary Figure 2. Colour overlay imaging highlighting vessels coloured according to vessel cross-sectional area on axial computer tomography images in the same 75-year-old male patient with idiopathic pulmonary fibrosis shown in supplementary figure 1. Emphysema is present in the upper lobes (A), with honeycombing evident in the right lower lobe and subtle reticulation seen in the left lung base (B). Vessels <5mm^2 in cross-sectional area (CSA)(red); Vessels <10mm^2 CSA (green); Vessels <15mm^2 CSA (yellow); Vessels <20mm^2 CSA (blue); Vessels >20mm^2 CSA (purple).

Dependent Variable	Independent Variable	Beta Coefficient	95% Confidence Interval	P value	R value
PVV	PVV5	2·85	2·70, 3.01	<0·0001	0·91
	PVV10	2·10	2·00, 2.20	<0·0001	0.93
	PVV15	1.79	1·71, 1.87	<0·0001	0.94
	PVV20	1·61	1.55, 1.68	<0·0001	0·95
	PVV>20	1·91	1·81, 2.00	<0·0001	0.92
	PVV>5	1.36	1.33, 1.39	<0·0001	0.98
DLco	PVV	-4·61	-5·42, -3·81	<0·0001	0·58
	PVV5	-14·71	-17·15,-12·26	<0·0001	0.60
	PVV10	-10·77	-12·51, -9·02	<0·0001	0.61
	PVV15	-5·64	-6·56, -4·72	<0·0001	0.61
	PVV20	-5·24	-6·10, -4.39	<0·0001	0·61
	PVV>20	-7·58	-9·42, -5·74	<0·0001	0.46
	PVV>5	-5.88	-7·04, -4.72	<0·0001	0.53

Supplementary Table 2. Linear correlations demonstrating the relationships between pulmonary vessel volume (PVV) and diffusing capacity for carbon monoxide (DLco) and various subdivisions of pulmonary vessel volume. The pulmonary vessel volume divisions are based on cross-sectional vessel area on axial computer tomography imaging. PVV5=vessels under 5mm^2 in cross-sectional area (CSA); PVV10=Vessels under 10mm^2 CSA; PVV15=Vessels under 15mm^2 CSA; PVV20=Vessels under 20mm^2 CSA; PVV>20=Vessels over 20mm^2 CSA; PVV>5=Vessels greater than 5 mm^2 CSA.

Pulmonary vessel volume subdivision	Hazard ratio (%)	P Value	95.0% Confidence Interval	
			Lower	Upper
PVV5	3·02	<0·0001	2·36	3.87
PVV10	2.29	<0·0001	1.91	2.75
PVV15	1·55	<0·0001	1.41	1·70
PVV20	1·51	<0·0001	1.38	1·65
PVV>20	2·14	<0·0001	1·83	2·49
PVV>5	1.75	<0·0001	1.57	1.96

Supplementary Table 3. Univariate Cox mortality analyses demonstrating mortality according to pulmonary vessel volume (PVV) subdivisions based on cross-sectional vessel area on axial computer tomography imaging. PVV5=vessels under 5mm^2 in cross-sectional area (CSA); PVV10=Vessels under 10mm^2 CSA; PVV15=Vessels under 15mm^2 CSA; PVV20=Vessels under 20mm^2 CSA; PVV>20=Vessels over 20mm^2 CSA. PVV>5=Vessels greater than 5 mm^2 CSA.

Variable (units are percentage unless specified)	Hazard ratio	P Value	95.0% Confidence Interval	
			Lower	Upper
CALIPER analysis				
PVV>5	1·79	<0·0001	1·60	1·99
Honeycombing	1·15	0·001	1·06	1·24
CALIPER PVV>5	1·37	0·0002	1·16	1·61
CALIPER Honeycombing	1·20	0·001	1·08	1·34
CPI	1·05	<0·0001	1·03	1·07
CALIPER PVV>5	1·58	<0·0001	1·37	1·82
CALIPER Honeycombing	1·24	<0·0001	1·11	1·38
GAP score (max score 8)	1.18	0.007	1.05	1.33

Supplementary Table 4. Multivariate Cox regression analysis demonstrating mortality separately for CALIPER indices (top white) and combined CALIPER, visually derived HRCT indices and pulmonary function tests (light grey). The GAP score was substituted for CPI in the final combined multivariate model but was a weaker predictor of mortality than the CPI (dark grey). PVV>5=pulmonary vessel volume of vessels greater than $5mm^2$ in cross-sectional area, CPI=composite physiologic index.

Composite score (Max score 3)	Hazard ratio	P Value	95% Confidence Interval		Harrells C Index
			Lower	Upper	
GAP index staging system	2.00	<0·0001	1·61	2.49	0.62
CALIPER-PVV>5-HC-CPI	2·36	<0·0001	1.94	2,87	0.67
CALIPER-PVV>5-HC	2·21	<0·0001	1.83	2.68	0.66
GAP index staging system	1·18	0.19	0.92	1.53	
CALIPER-PVV>5-HC-CPI	2.15	<0·0001	1.68	2.74	
GAP index staging system[b]		0·17	-0.07	0.43	
CALIPER-PVV>5-HC-CPI[b]		0·001	0.55	1.00	
GAP index staging system	1.44	0·002	1.14	1.82	
CALIPER-PVV>5-HC	1.92	<0·0001	1.55	2.37	
GAP index staging system[b]		0·001	0.15	0.59	
CALIPER-PVV>5-HC[b]		0·001	0.45	0.88	

Supplementary Table 5. Univariate Cox regression analyses (white) comparing the Gender, Age, Physiology (GAP) index staging system with scores derived from two separate hazard ratio formulae. The first formula derived from the hazard ratios of the independent predictors of mortality in a combined visual, CALIPER and pulmonary function index multivariate model (CALIPER-PVV>5-HC-CPI). The second formula derived from the hazard ratios of CALIPER variables that were independent predictors of mortality (CALIPER-PVV>5-HC). Bivariate analyses (light and dark grey) compared the hazard ratio derived scores to the GAP index staging system with the results confirmed on bootstrapping of 1000 samples. CPI=Composite Physiologic Index, [b]=bootstrapped results.

REFERENCES

1.   Quanjer PH. Standardized lung function testing. *Eur Respir J - Suppl*. 1993;**6**:1-100.
2.   Wells AU, Desai SR, Rubens MB, et al. Idiopathic pulmonary fibrosis: a composite physiologic index derived from disease extent observed by computed tomography. *Am J Respir Crit Care Med*. 2003;**167**:962-9.
3.   Matsuoka S, Washko GR, Yamashiro T, et al. Pulmonary hypertension and computed tomography measurement of small pulmonary vessels in severe emphysema. *AmJRespirCrit Care Med*. 2010;**181**(3):218-25.
4.   Estépar RSJ, Kinney GL, Black-Shinn JL, et al. Computed Tomographic Measures of Pulmonary Vascular Morphology in Smokers and Their Clinical Implications. *Am J Respir Crit Care Med*. 2013;**188**(2):231-9.