EXTENDED REPORT

Evidence-based detection of pulmonary arterial hypertension in systemic sclerosis: the DETECT study

J Gerry Coghlan,[1] Christopher P Denton,[2] Ekkehard Grünig,[3] Diana Bonderman,[4] Oliver Distler,[5] Dinesh Khanna,[6] Ulf Müller-Ladner,[7] Janet E Pope,[8] Madelon C Vonk,[9] Martin Doelberg,[10] Harbajan Chadha-Boreham,[11] Harald Heinzl,[12] Daniel M Rosenberg,[11] Vallerie V McLaughlin,[6] James R Seibold,[13] on behalf of the DETECT study group

Handling editor Tore K Kvien

▶ Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/annrheumdis-2013-203301).

For numbered affiliations see end of article.

Correspondence to
Dr J Gerry Coghlan, Cardiology Department, Royal Free Hospital, Pond Street, Hampstead, London NW3 2QG, UK; gerry.coghlan@nhs.net

JGC, CPD and EK contributed equally.

Accepted 14 April 2013
Published Online First 18 May 2013

To cite: Coghlan JG, Denton CP, Grünig E, *et al*. *Ann Rheum Dis* 2014;**73**:1340–1349.

ABSTRACT

Objective Earlier detection of pulmonary arterial hypertension (PAH), a leading cause of death in systemic sclerosis (SSc), facilitates earlier treatment. The objective of this study was to develop the first evidence-based detection algorithm for PAH in SSc.

Methods In this cross-sectional, international study conducted in 62 experienced centres from North America, Europe and Asia, adults with SSc at increased risk of PAH (SSc for >3 years and predicted pulmonary diffusing capacity for carbon monoxide <60%) underwent a broad panel of non-invasive assessments followed by diagnostic right heart catheterisation (RHC). Univariable and multivariable analyses selected the best discriminatory variables for identifying PAH. After assessment for clinical plausibility and feasibility, these were incorporated into a two-step, internally validated detection algorithm. Nomograms for clinical practice use were developed.

Results Of 466 SSc patients at increased risk of PAH, 87 (19%) had RHC-confirmed PAH. PAH was mild (64% in WHO functional class I/II). Six simple assessments in Step 1 of the algorithm determined referral to echocardiography. In Step 2, the Step 1 prediction score and two echocardiographic variables determined referral to RHC. The DETECT algorithm recommended RHC in 62% of patients (referral rate) and missed 4% of PAH patients (false negatives). By comparison, applying European Society of Cardiology/European Respiratory Society guidelines to these patients, 29% of diagnoses were missed while requiring an RHC referral rate of 40%.

Conclusions The novel, evidence-based DETECT algorithm for PAH detection in SSc is a sensitive, non-invasive tool which minimises missed diagnoses, identifies milder disease and addresses resource usage.

INTRODUCTION

The diagnosis of pulmonary arterial hypertension (PAH) is defined at right heart catheterisation (RHC) by a mean pulmonary arterial pressure (mPAP) of ≥25 mm Hg with a pulmonary capillary wedge pressure (PCWP) of ≤15 mm Hg.[1] Additional diagnostic criteria may include a normal or reduced cardiac output[1] or a pulmonary vascular resistance (PVR) of >3 Wood units.[2] PAH includes diverse clinical phenotypes, prominent among which is systemic sclerosis (SSc, scleroderma) where PAH has emerged as a leading cause of death.[3 4] Three-year survival for SSc patients with PAH has been estimated to be 56% compared with 94% in those without PAH.[5] Observational studies have demonstrated that mortality remains high in SSc patients with PAH despite current best therapy.[6 7]

Poor outcome of PAH in SSc may be partially explained by disease-related comorbidities but also by delay in diagnosis. One recent study observed a better prognosis in subjects identified in an active screening programme compared with those identified in the course of routine practice,[8] suggesting potential benefit of intervention earlier in the course of disease. This is consistent with the beneficial treatment effects demonstrated in early PAH.[9] Current screening recommendations are largely based on consensus.[1 2] Several organisations, including the American College of Cardiology Foundation/American Heart Association and the European Society of Cardiology/European Respiratory Society (ESC/ERS), have published a variety of screening recommendations relying mainly on symptoms and abnormal findings on transthoracic echocardiography.[1 2 10] Other clinical tools include N-terminal probrain natriuretic peptide (NTproBNP) as a marker of myocardial stress,[11] and disproportionately reduced pulmonary diffusing capacity for carbon monoxide (DLCO).[12–14] The most widely used echocardiographic parameter, tricuspid regurgitant jet (TR) velocity, does not accurately reflect invasive pressures and is not present in all patients.[15 16] Furthermore, while TR velocity recommendations are very specific in current guidelines, recommendations regarding other evidence of PAH (eg, symptoms) are less detailed; thus, application is likely to be variable between clinicians. More importantly, no previous screening studies have systematically performed RHC in all patients, precluding assessment of the rate of missed diagnoses (false negatives).

Our study provides evidence-based data guided by the principles of screening[17]: (A) employing rigorous methodology using the appropriate cross-sectional study design in order to determine the performance characteristics of the screening algorithm (sensitivity, specificity, etc); (B) evaluating accessible and feasible real-world screening tools; (C) identifying patients during an asymptomatic

phase of disease (the SSc population in this study has symptoms which are not PAH specific) and (D) identifying patients for whom subsequent management is appropriate. Using the anchor of systematic RHC, the objective of the DETECT study was to develop a detection algorithm for PAH in SSc patients that would minimise the number of missed PAH diagnoses, while optimising the use of diagnostic RHC.

METHODS

Study design

DETECT was designed as a cross-sectional study in which RHC and echocardiography were systematically conducted according to standardised procedures. Serum laboratory testing and data management were performed centrally, and data quality was rigorously monitored. DETECT was conducted in accordance with the Declaration of Helsinki and its amendments, followed the International Conference on Harmonisation Guideline for Good Clinical Practice, and was approved by local institutional review boards/ethics committees. All patients provided written informed consent.

Study population

Sixty-two experienced centres (managing at least 40 SSc patients) from 18 countries in North America, Europe and Asia participated in the study between 2008 and 2011. Patients aged ≥18 years with a diagnosis of SSc (American College of Rheumatology criteria,[18] including patients with other connective tissue diseases who met these criteria) of >3 years' duration from first non-Raynaud's symptom and a predicted DLCO of <60% (to enrich for a higher likelihood of PAH), were included. Patients were excluded if they had pulmonary hypertension (PH) confirmed by RHC prior to enrolment, were receiving recognised advanced PH therapy, had a forced vital capacity (FVC) <40% of predicted, renal insufficiency, previous evidence of clinically relevant left heart disease, or were pregnant.

Patients were classified as either non-PH, or WHO group 1 PH (PAH), WHO group 2 PH (PH due to left heart disease), or WHO group 3 PH (PH due to lung disease/hypoxia), according to current guidelines.[10 19] WHO group 3 definition was based on Study Scientific Committee consensus applying a conservative interpretation of the literature to minimise misclassification of patients with evident lung disease as PAH. Classification definitions are summarised in figure 1.

Data collection and analysis

A broad range of variables potentially associated with PAH in SSc were assessed (112 in total; see online supplementary text). Four groups of variables were: (A) standard demographic and clinical parameters (68 variables in total, eg, SSc disease duration from first non-Raynaud's symptom, SSc subtype, SSc symptoms and organ involvement, general medical history, standard physical examination, 6-min walk distance, standard pulmonary function tests); (B) serum tests analysed by a central laboratory (antinuclear antibody profile (five antibodies), NTproBNP, endothelin-1, von Willebrand factor antigen, C-reactive protein, serum urate, creatinine, erythrocyte sedimentation rate, estimated glomerular filtration rate); (C) electrocardiography (ECG; right ventricular strain, right axis deviation, right bundle branch block) and (D) echocardiography according to standardised procedures (28 variables in total, eg, right atrium (RA) area, right ventricle (RV) area, RA diameter, TR velocity, tricuspid annular plane systolic excursion). To minimise bias, RHC as the confirmatory diagnostic test (conducted according to standardised procedures), was performed in all patients following collection of aforementioned data. Serious adverse events related to any study-mandated procedure (eg, RHC) were collected.

Statistical methodology

It was planned to enrol approximately 500 SSc patients including the planned number of 70 patients testing positive for PAH. This planned sample size took into consideration feasibility aspects of the study and assumed a prevalence rate of 14%.[11] This sample size was calculated to allow an estimation of 90% sensitivity of the detection algorithm with a precision of ±7.5%. At a similar level of expected specificity, its precision is superior, due to the higher prevalence of non-PAH.

PAH and non-PH groups were described using summary statistics; sample size, mean, SD, median, upper and lower quartiles, minimum and maximum, and 95% CIs of the mean and median for quantitative data and frequencies (counts and percentages) for qualitative and categorical data.

Logistic regression modelling was the main analytical method, including linear and non-linear functional relations, where the binary outcome variable was PAH versus non-PH. Model-building entailed use of statistical procedures; variable selection was informed by clinical judgement and internal validation of models was performed via the bootstrap method.

Statistical analysis for selecting predictive variables and developing the detection algorithm for risk prediction of PAH was performed stepwise in three broad stages (see online supplementary text, tables S2–S5 and figures S1 and S2): (A) univariable and multivariable logistic regression models with RHC-based classification of PAH outcome, were applied within each of the four above-mentioned groups of candidate variables to select those associated with PAH; (B) the selected variables were further reduced across groups by using multivariable logistic regression; using nominal group technique, the Study Scientific Committee excluded some variables based on lack of clinical plausibility and/or feasibility with particular regard to resource limitations in standard practice and (C) a two-step decision tree was constructed based on two multivariable logistic regression models. The first step (sensitivity set at 97%) of the decision tree included non-echocardiographic tests to produce a risk prediction score that allowed exclusion of patients at low risk of having PAH and determined referral to echocardiography for the other patients. In the second step (specificity set at 35%), the risk score from Step 1 was combined with echocardiographic tests to produce the final PAH risk prediction score to determine if a patient should be referred to RHC for diagnosis. Spline functions with three knots were used in the models to adequately address non-linear relationships, which were initially identified by quadratic functions during the model-building process. Discriminatory performance to distinguish between PAH and non-PH patients was examined by receiver operating characteristic (ROC) curve analysis. The ROC area under the curve (AUC) formed the criterion for assessing the discriminatory ability of a model. Nomograms[20] were derived from the two multivariable risk prediction models (see online supplementary text) to allow classification of patients into risk sets for referral to echocardiography (Step 1) and RHC (Step 2). An alternative algorithm with 65% specificity set in the second step was also evaluated, as was the application of the ESC/ERS guidelines to the DETECT population.[1] The performance measures of the decision tree and its internal validation using bootstrap methodology are described in detail in the online supplementary text and tables S8–S11.

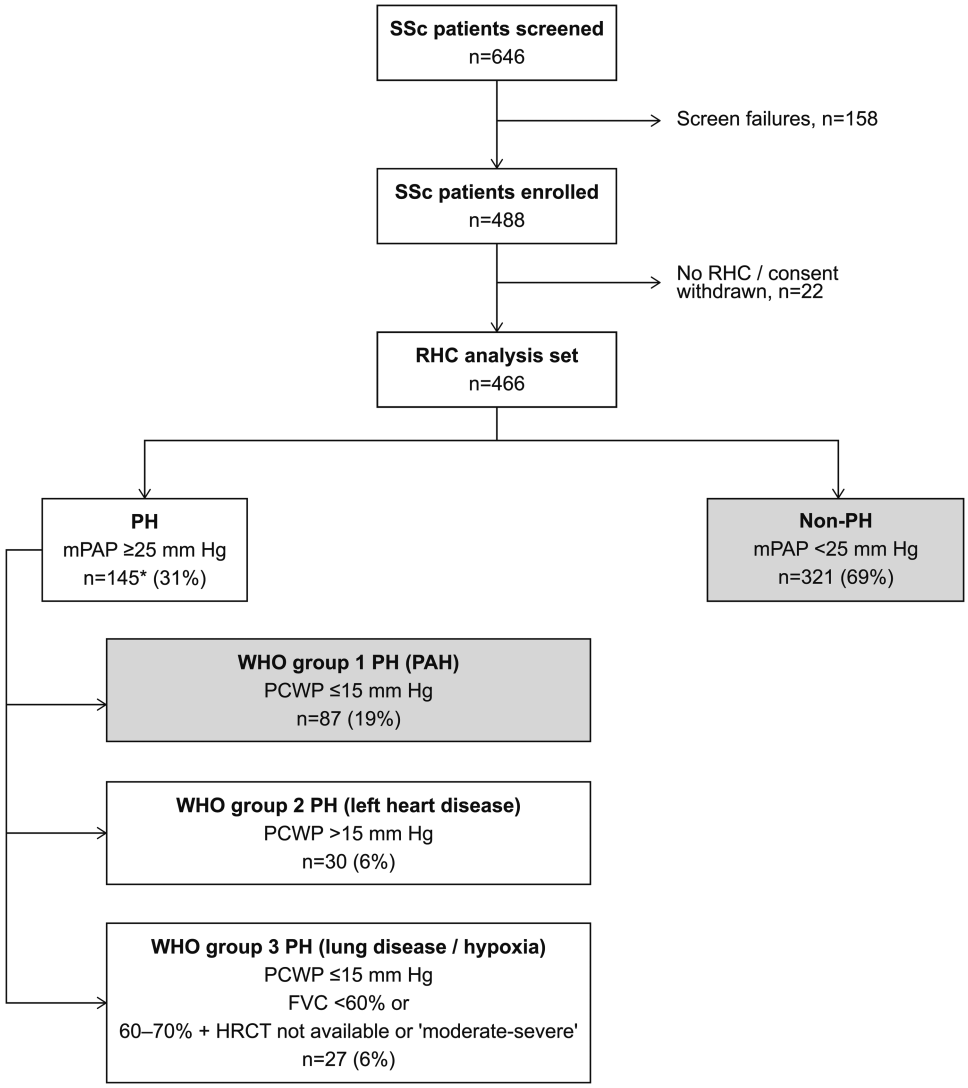


Figure 1 Patient disposition. The results reported here focus on the 408 SSc patients with PAH (n=87) and those without PH (n=321; grey boxes). *One patient could not be assigned to a PH group due to a missing PCWP value. FVC, forced vital capacity; HRCT, high-resolution CT; mPAP, mean pulmonary arterial pressure; PAH, pulmonary arterial hypertension; PH, pulmonary hypertension; PCWP, pulmonary capillary wedge pressure; RHC, right heart catheterisation; SSc, systemic sclerosis.

RESULTS

Of 646 SSc patients screened, 158 did not meet the eligibility criteria (mostly due to DLCO values ≥60% of predicted). Of the 488 patients enrolled, 466 underwent RHC which revealed PH in 31% of patients (n=145) and PAH (WHO group 1 PH) in 19% (n=87; figure 1). Our results focus on these 87 PAH patients versus the 321 non-PH patients.

Patient characteristics, including RHC findings, are summarised in table 1.

In most patients (64%), PAH was mild (WHO functional class I or II), with moderately elevated mPAP and pulmonary vascular resistance and preserved mean cardiac index. In this population, among other variables, exercise capacity on 6-min walk test and dyspnoea were not associated with the presence of PAH. However, compared with non-PH patients, PAH patients were older, more likely to be male, in higher (more severe) WHO functional class, more likely to have the limited cutaneous form of SSc, to be anticentromere antibody (ACA) positive and to have a history of telangiectasias, had worse gas transfer (as assessed by DLCO), higher serum urate and NTproBNP levels, were more likely to have a right ventricular strain and right axis deviation on ECG, had larger RA and RV areas and higher TR velocity. However, when analysing commonly advocated TR velocity thresholds for PAH suspicion,[1 21] 20% of PAH patients were found to have a TR velocity of <2.5 m/s, 36% had a TR velocity of ≤2.8 m/s, and 63% had a TR velocity of ≤3.4 m/s (including 7% of PAH patients with undetectable TR velocity). Within the total DETECT cohort, 49% had a TR velocity of <2.5 m/s (including 13% with undetectable TR velocity). Several other echocardiographic variables (eg, tricuspid annular plane systolic excursion) were associated with the presence of PAH but did not progress to the final model (see below).

Following univariable and multivariable analyses, and clinical judgment of the Study Scientific Committee (based on feasibility and clinical plausibility), from an initial 112 variables, 13 were selected based on their discriminatory ability to detect PAH (see online supplementary text and table S3). These formed the basis for constructing a detection algorithm. To align the algorithm with real-world practice where the rheumatologist accesses

Table 1 Patient characteristics

	Non-PH group (N=321)	PAH group (N=87)
Demographics		
Male, n/N (%)	53/320 (16.6)	22/86 (25.6)
Age (years), mean (SD)	54.7 (11.8)	61.1 (9.8)
Caucasian, n/N (%)	258/319 (80.9)	81/87 (93.1)
Body mass index (kg/m^2), n	317	86
Mean (SD)	25.2 (5.6)	26.1 (5.5)
SSc characteristics		
SSc duration* (months), n	319	87
Mean (SD)	130.2 (96.1)	163.0 (130.3)
SSc subtype, n/N (%)		
Diffuse	115/315 (36.5)	18/86 (20.9)
Limited	171/315 (54.3)	61/86 (70.9)
Mixed/overlap	29/315 (9.2)	7/86 (8.1)
Current/past telangiectasias, n/N (%)	218/321 (67.9)	76/87 (87.4)
Functional capacity		
6-min walk distance (m), n	243	66
Mean (SD)	412.5 (107.2)	389.7 (106.6)
Borg dyspnoea index, n	240	66
Mean (SD)	2.6 (1.8)	3.1 (2.1)
WHO functional class, n	306	87
I/II/III/IV, n (%)	133 (43.5)/123 (40.2)/50 (16.3)/0	16 (18.4)/40 (46.0)/30 (34.5)/1 (1.1)
Pulmonary function tests, n	321	87
DLCO % predicted		
Mean (SD)	48.0 (9.2)	43.3 (10.5)
FVC % predicted/DLCO % predicted		
Mean (SD)	1.8 (0.5)	2.2 (0.7)
Serum markers, n	306	80
NTproBNP (pg/ml)		
Mean (SD)	230.0 (538.6)	516.4 (805.0)
Serum urate (mg/100 ml)		
Mean (SD)	4.7 (1.5)	5.9 (1.5)
ACA positive, n/N (%)	77/306 (25.2)	40/80 (50.0)
Electrocardiography, n/N (%)		
Right ventricular strain present	7/291 (2.4)	12/83 (14.5)
Right axis deviation† present	10/291 (3.4)	11/83 (13.3)
Echocardiography		
Right atrium area (cm^2), n	286	82
Mean (SD)	13.4 (4.7)	17.1 (6.2)
Right ventricle area (cm^2), n	291	82
Mean (SD)	15.0 (5.4)	19.3 (6.8)
TR velocity (m/s), n	255	78
Mean (SD)	2.4 (0.5)	3.1 (0.7)
TR velocity (m/s), n/N (%)		
'No TR' ticked	48/303 (15.8)	6/84 (7.1)
≤2.8	214/303 (70.6)	30/84 (35.7)
>2.8 to ≤3.4	37/303 (12.2)	23/84 (27.4)
>3.4	4/303 (1.3)	25/84 (29.8)
Haemodynamics on right heart catheterisation		
mPAP (mm Hg), n	321	87
Mean (SD)	17.6 (3.8)	32.5 (8.3)
PCWP (mm Hg), n	318	87
Mean (SD)	8.5 (3.6)	10.3 (3.2)
PVR (dyn·sec/cm^5), n	318	87
Mean (SD)	145.4 (64.6)	370.6 (225.8)
Cardiac index (l/min/m^2), n	317	86
Mean (SD)	3.0 (0.7)	2.9 (0.6)

*From date of first non-Raynaud's symptom.
†QRS axis ≥90°.
ACA, anticentromere antibody; DLCO, pulmonary diffusing capacity for carbon monoxide; FVC, forced vital capacity; mPAP, mean pulmonary arterial pressure; NTproBNP, N-terminal probrain natriuretic peptide; PAH, pulmonary arterial hypertension; PH, pulmonary hypertension; PCWP, pulmonary capillary wedge pressure; PVR, pulmonary vascular resistance; SSc, systemic sclerosis; TR, tricuspid regurgitant jet.

Table 2 Logistic regression models

	Univariable logistic regression models*	Multivariable logistic regression models (two-step decision tree)		
	ROC AUC,% (95% CI)	Estimated coefficient (95% CI)	p value	ROC AUC,% (95% CI)
Step 1				84.4 (79.5 to 89.8)
Intercept		−12.488 (−16.372 to −8.603)		
FVC % pred./DLCO % pred.	71.5 (65.6 to 77.4)	1.149 (0.566 to 1.731)	<0.001	
Current/past telangiectasias	59.7 (55.4 to 64.1)	1.156 (0.336 to 1.975)	0.006	
Serum ACA (presence)	62.4 (56.4 to 68.4)	0.753 (0.133 to 1.373)	0.017	
Serum NTproBNP (log_{10})	67.5 (60.9 to 74.2)	0.915 (0.308 to 1.521)	0.003	
Serum urate	71.9 (65.9 to 77.9)	1.247 (0.497 to 1.997)	<0.001	
Serum urate (spline component†)		−1.132 (−2.048 to −0.215)	–	
Right axis deviation (presence)	54.9 (51.1 to 58.7)	1.850 (0.507 to 3.193)	0.007	
Step 2				88.1 (82.4 to 92.3)
Intercept		−2.452 (−5.747 to 0.844)		
Linear predictor Step 1		0.891 (0.559 to 1.224)	<0.001	
Right atrium area‡	71.2 (65.0 to 77.3)	0.075 (−0.004 to 0.154)	0.062	
TR velocity	79.5 (73.7 to 85.3)	0.209 (−1.117 to 1.534)	<0.001	
TR velocity (spline component†)		2.656 (0.380 to 4.933)	–	

*The ROC AUC values of the corresponding univariable logistic regression models are added for the sake of comparison.
†Serum urate and TR velocity were included in the respective models using restricted cubic splines with three knots. Knots for serum urate were selected at 3.3, 4.7 and 7.1 mg/100 ml and for TR velocity were selected at 2, 2.5 and 3.4 m/s. For each of these two variables only one p value is presented indicating its overall effect.
‡When right ventricle area was used instead of right atrium area, the p value was 0.035; the rest of the p values remained with similar results.
ACA, anticentromere antibody; DLCO, pulmonary diffusing capacity for carbon monoxide; FVC, forced vital capacity; NTproBNP, N-terminal probrain natriuretic peptide; ROC AUC, area under the receiver operating characteristic curve; TR, tricuspid regurgitant jet.

non-echocardiographic data prior to referral to a cardiologist for echocardiography, the 13 variables were divided into nine non-echocardiographic variables (candidate variables in Step 1 of the algorithm) and four echocardiographic variables (candidate variables in Step 2). Subsequent multivariable analysis (stepwise forward procedure) resulted in six simple assessments being included in Step 1 of the algorithm to determine the need for referral to echocardiography (table 2). These were: FVC % predicted/DLCO % predicted, current/past telangiectasias, serum ACA, serum NTproBNP, serum urate and right axis deviation on ECG. Two echocardiographic variables (RA area and TR velocity) were included in Step 2 in order to determine the need for referral for RHC. Additionally, Step 2 included the carried-forward Step 1 risk prediction score (risk points) (table 2). Of all statistically selected final variables, the Study Scientific Committee replaced one echocardiographic variable (RV area) with another (RA area), because the latter is regarded as easier to assess and likely to be more reproducible. This replacement had a minimal effect on the performance of the Step 2 model (AUC 89% and 88% using RV area and RA area, respectively). The multivariable logistic regression models for the two-step decision tree are summarised in table 2.

Sensitivity (97% in Step 1) and specificity (35% in Step 2) were selected by the Study Scientific Committee with the aim of minimising the number of missed PAH diagnoses. The performance of the resulting DETECT algorithm is presented in figure 2, and a nomogram of the DETECT algorithm for use in clinical practice is shown in figure 3. Internal bootstrap validation generated consistent results and confirmed overall performance of the algorithm (see online supplementary text and tables S8 and S9). Exclusion of any single variable from the DETECT algorithm had only a small impact on model performance (see online supplementary text and table S11). If more than one variable is missing, the model cannot be used reliably; in clinical practice, a single missing variable should be handled as described in the legend to figure 3. Depending on the risk points in Step 1 of the algorithm, Step 2 (echocardiography) may not be needed for recommending RHC in some patients (figure 3).

As presented in table 3, the rate of missed PAH diagnoses was 4% (n=3) applying the DETECT algorithm. The three missed PAH patients had mPAP values of 26, 25 and 30 mm Hg, PCWP values of 11–12 mm Hg, PVR values of 238–257 dyn·sec/cm^5, TR velocities of 2.9, 2.5 and 2.6 m/s, RA areas of 16.4, 15.0 and 8.5 cm^2, NTproBNP levels of 204, 47 and 71 pg/ml, no right axis deviation on ECG, and were ACA negative. The 4% missed diagnoses rate of the DETECT algorithm compares with 29% (n=24; see patient characteristics in online supplementary table S7) based on current ESC/ERS guidelines.[1] The proportion of RHC which did not confirm a diagnosis of PAH was similar between the DETECT algorithm and the ESC/ERS guidelines (65% vs 60%). Applying the DETECT algorithm recommended referral of 62% of patients to RHC. Reducing the RHC referral rate from 62% to 41% (ie, to a similar level as the 40% RHC referral rate observed with the ESC/ERS guidelines) increased the rate of missed PAH diagnoses to 15% (n=11) which is still lower than the 29% achieved with the ESC/ERS guidelines (table 3).

Among the 466 patients who underwent RHC, one patient had a haematoma caused by accidental carotid puncture. This was managed without hospital admission or transfusion.

DISCUSSION

DETECT was a large, multicentre, real-world, cross-sectional study with detailed population characterisation, standardised RHC and echocardiography procedures, central serum testing, central data management and rigorous data monitoring. It is the first PAH detection study to undertake systematic RHC in all patients and to develop an evidence-based algorithm using simple clinical data and non-invasive tests for earlier identification of PAH in a mildly symptomatic population. The DETECT study demonstrates that within this cohort of SSc patients, PAH is much more common than previous studies have suggested,[21 22]

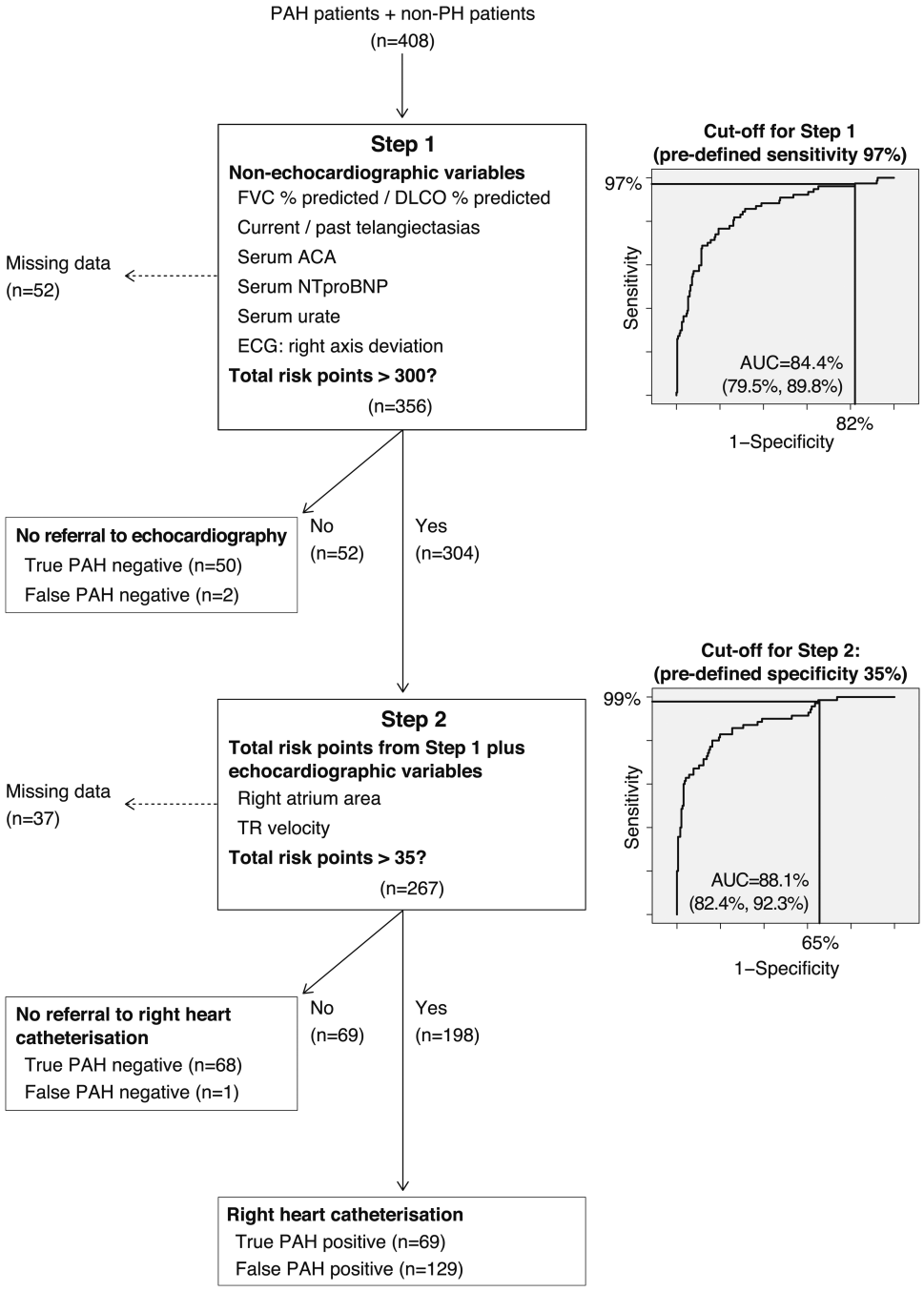


Figure 2 Two-step decision tree for detection of pulmonary arterial hypertension in systemic sclerosis patients: the DETECT algorithm. Of the 408 SSc patients (87 PAH and 321 non-PH) at risk for PAH (SSc of >3 years' duration, DLCO <60% of predicted, FVC ≥40% of predicted), data from 319 patients (72 PAH and 247 non-PH) were used for construction of the algorithm. All patients underwent right heart catheterisation. Sensitivity and specificity of the two steps of the algorithm (and the corresponding risk point cut-offs) were selected by the Study Scientific Committee with the aim of minimising the number of missed PAH diagnoses. Step 1: A complete dataset was available for 356 patients. The combined discriminatory ability of the six selected non-echocardiographic variables expressed as the AUC of the ROC curve was 84.4% (95% CI 79.5% to 89.8%) showing good discriminatory performance and no statistically significant lack of fit (see online supplementary appendix 5). At Step 1, a predefined sensitivity cut-off of 97% (corresponding to >300 risk points, compare figure 3), determined no referral to echocardiography in 52 patients. Among these, 50 were true negatives (patients without PAH on right heart catheterisation) and two were false negatives (PAH confirmed on right heart catheterisation). Step 2: A complete dataset was available for 267 patients. The AUC of the ROC curve for the total risk points from Step 1, plus the two selected echocardiographic variables, was 88.1% (95% CI 82.4% to 92.3%). A predefined specificity cut-off of 35% (corresponding to >35 risk points, compare figure 3), determined no referral to right heart catheterisation in 69 patients. Among these, 68 were true negatives and one was a false negative. Right heart catheterisation in the remaining 198 patients yielded 69 true positives (PAH confirmed) and 129 false positives. Thus, overall, the algorithm missed 3 (4%) out of the 72 PAH patients who had sufficient data to be included in the analysis. Note that the algorithm uses cut-offs for the risk points of the two steps only but not for individual parameters. ACA, anticentromere antibody; AUC, area under the curve; DLCO, pulmonary diffusing capacity for carbon monoxide; FVC, forced vital capacity; NTproBNP, N-terminal probrain natriuretic peptide; PAH, pulmonary arterial hypertension; PH, pulmonary hypertension; ROC, receiver operating characteristic; SSc, systemic sclerosis; TR, tricuspid regurgitant jet.

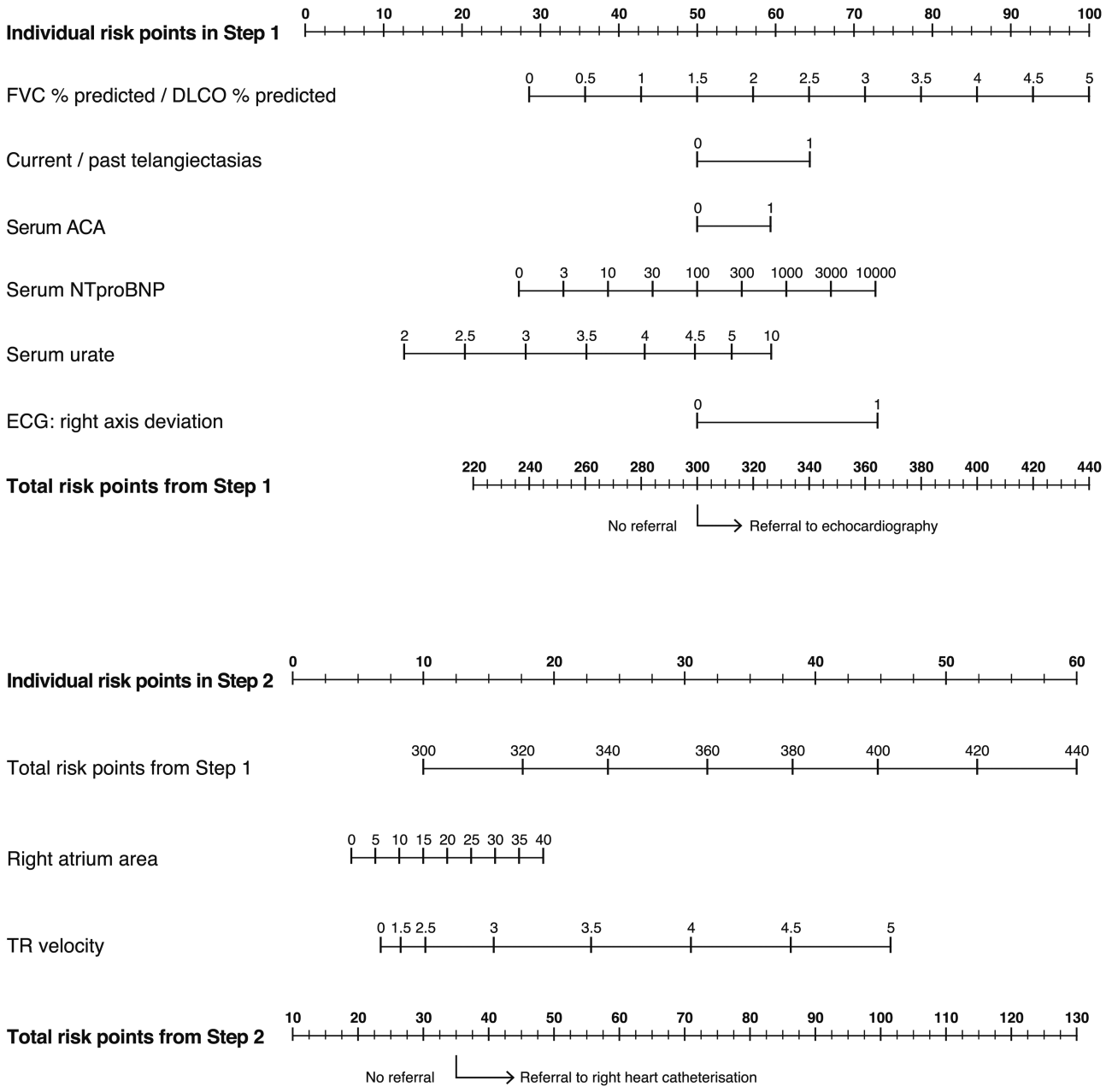


Figure 3 Nomograms for practical application of the DETECT algorithm: determination of the likelihood of pulmonary arterial hypertension and cut-off points for decision to refer a patient to echocardiography (Step 1) and subsequent right heart catheterisation (Step 2). At Step 1 (top panel), risk points for each of the six non-echocardiographic variables are calculated by reading from 'Individual risk points in Step 1' and adding them up to obtain a total. If the 'Total risk points from Step 1' is >300 (corresponding to a sensitivity of 97% as selected by the Study Scientific Committee) the patient is referred to echocardiography. Similarly, at Step 2 (bottom panel), risk points for the carried forward 'Total risk points from Step 1' and the two echocardiographic variables are calculated by reading from the 'Individual risk points in Step 2'. If the 'Total risk points from Step 2' is >35 (corresponding to a specificity of 35% as selected by the Study Scientific Committee) the patient is referred to right heart catheterisation. Alternatively, being less conservative (65% predefined specificity at Step 2), the patient would be referred to right heart catheterisation if 'Total risk points from Step 2' is >40 (compare table 3 for the performance of these two options). Note that all variables will always contribute risk points irrespective of the measured value; for example, a negative serum ACA will contribute 50 risk points. Exclusion of any single variable from the DETECT algorithm has only a small impact on model performance (see online supplementary appendix 9). If a single Step 1 variable is missing it should be assigned 50 risk points, with the exception of current/past telangiectasias which should be assigned 65 points. If a single Step 2 variable is missing it should be assigned 10 points. The nomograms cannot be reliably used if more than one variable out of the eight total variables is missing. ACA, anticentromere antibody; DLCO, pulmonary diffusing capacity for carbon monoxide; FVC, forced vital capacity; NTproBNP, N-terminal probrain natriuretic peptide; TR, tricuspid regurgitant jet.

and can be identified before symptoms are sufficiently advanced to be discriminated from general SSc symptoms.

Previous screening initiatives[21 22] without systematic RHC in all patients, resulted in guideline recommendations based on symptoms and echocardiography without information on the number of missed PAH diagnoses (false negatives).[1 2 10] Consequently, PAH continues to be diagnosed late with advanced symptoms.[6 8 23] Given the evidence that early intervention may delay morbidity in PAH,[9] and that screening programmes that allow earlier treatment in patients with PAH

Table 3 Model performance: comparison of PAH detection approaches

Approach	RHC referral rate, % (positive detection assessments/all patients)	Overall missed PAH diagnoses, % (false negatives)	Overall sensitivity, %	Overall specificity, %	Overall PPV, %	Overall NPV, %
Primary analysis						
DETECT algorithm N=319	62	4	96	48	35	98
Other analyses						
DETECT algorithm with 65% specificity at Step 2 N=319	41	15	85	72	47	94
ESC/ERS guidelines*[1] N=371	40	29	71	69	40	89

*Evaluated on a subset of patients (N=371) with available data for the variables defined in the guideline, using the following criteria for RHC referral[1]: (a) Tricuspid regurgitant jet velocity >3.4 m/s; or (b) Tricuspid regurgitant jet velocity >2.8–≤3.4 m/s AND symptomatic (defined as at least one of the following DETECT parameters: current anginal pain, current syncope/near syncope, current dyspnoea, presence of peripheral oedema); or (c) Tricuspid regurgitant jet velocity ≤2.8 m/s AND symptomatic (defined as above) AND presence of additional echocardiography variables suggestive of pulmonary hypertension (defined as right atrium area >16 cm^2 and/or ratio of right ventricular diameter/left ventricular end diastolic diameter >0.8).
ESC/ERS, European Society of Cardiology/European Respiratory Society; NPV, negative predictive value; PAH, pulmonary arterial hypertension; PPV, positive predictive value (confirmed PAH out of all RHC referrals); RHC, right heart catheterisation.

related to SSc may change the way patients are managed and may thus improve prognosis,[8 24] the DETECT algorithm addresses a major medical need in this patient population. Among all patient populations at risk for developing PAH, SSc appears most suitable for screening programmes in terms of prevalence and feasibility.

To align the DETECT approach with clinical practice where the rheumatologist may have non-echocardiographic data available prior to referral to a cardiologist for echocardiography, a corresponding two-step algorithm was developed. This will limit echocardiography referral to patients at increased PAH risk. Of course, the algorithm can also be used if all data including echocardiography are available at the same time. Sensitivity and specificity of the two steps of the algorithm (and the corresponding risk point cut-offs) were selected by the Study Scientific Committee with the aim of minimising the number of missed PAH diagnoses as compared with the ESC/ERS guidelines[1] (table 3). Given that the high overall sensitivity (96%) selected for the DETECT algorithm is associated with reduced specificity (48%), the rate of RHC required in the high-risk SSc population included in DETECT is substantial (62%). However, the proportion of RHC performed that did not confirm a diagnosis of PAH was similar between the DETECT algorithm and the ESC/ERS guidelines (65% vs 60%). Furthermore, overall RHC usage when applying the DETECT algorithm in clinical practice is likely to be lower than when compared with current guideline recommendations, since the latter are applied to the general SSc population, without enrichment for high PAH risk. For pragmatic reasons (sample size, reasonable expected positive predictive value) and ethical considerations (mandated RHC in all patients) the DETECT algorithm is designed for application in a high-risk SSc population (inclusion criteria DLCO <60% and SSc disease duration >3 years). The prevalence of PAH in SSc patients with a DLCO ≥60% is not known but may be as low as 1.2%, sevenfold lower than in patients with a DLCO <60%.[21] Screening these patients would not reduce the rate of missed PAH diagnoses (4% in DETECT) much further, but would increase the rate of false positives; the number of RHCs needed per PAH diagnosis was three in DETECT, but would be six in a population with 10% prevalence and 11 in a population with 5% prevalence. In a regularly screened unselected SSc population, incidence and prevalence of PAH were found to increase progressively, with onset generally after 3 years of disease.[25] Based on these data, it is likely that few PAH patients were missed in DETECT as a result of the required minimum SSc disease duration, and that this criterion contributed to selection of a high-risk population. It is possible, however, that some PAH patients with very early SSc and preserved DLCO were missed.

Adjusting the specificity of the DETECT algorithm to recommend similar RHC referral rates (in the DETECT high-risk population) as the ESC/ERS guidelines still substantially reduces the number of missed PAH diagnoses (table 3). As shown in this study and in previous published experience,[26] RHC is a safe technique in experienced centres.

Clinical plausibility, feasibility and applicability of the final selected variables were assured by expert input, and their robustness was internally validated. Some parameters previously identified as predictive of PAH were confirmed in DETECT, such as FVC/DLCO,[27] telangiectasias,[27] ACA,[27 28] NTproBNP,[11 29] right axis deviation on ECG[30] and TR velocity.[8 22] Serum urate has not been described previously as being predictive of PAH, but was identified as such in this study where low values were associated with a low PAH risk. There is, however, some support for an association in the literature; in a study of 228 patients, serum urate levels were significantly higher in those with PAH than in age-matched controls,[31] findings that have been corroborated elsewhere.[32] Interestingly, we have demonstrated the limited utility of the two main components of current guidelines, that is, symptoms and echocardiography: dyspnoea, a prominent symptom of PAH, did not discriminate between PAH and absence of PH (which is consistent with its lack of sensitivity in identifying cardiopulmonary compromise in early disease perhaps due to SSc-associated restricted musculoskeletal mobility), and TR velocity alone would have missed 20% of PAH patients when using a PAH suspicion threshold of ≥2.5 m/s, 36% when using a threshold of >2.8 m/s and 63% when using a threshold of >3.4 m/s.

An expert consensus on criteria for referring SSc patients to RHC has been published recently.[33] The objectives, methodology and population of this study were different from those of DETECT (consensus-based assessment of symptomatic patients with suspected PH rather than prospective data-driven assessment of a primarily screening population without requirement for PAH suspicion). Of the criteria proposed by the expert

consensus, dyspnoea, physical findings related to the right heart and WHO functional class were not sufficiently predictive to proceed to the final DETECT algorithm. DLCO (and DLCO corrected for alveolar volume) was less predictive than FVC/DLCO. DETECT confirmed the value of TR velocity as one component of the algorithm despite poor performance as a single parameter. The predictive value of both RV and RA dilation was confirmed in DETECT; RA area is part of the final algorithm.

The inclusion criteria selected for prevalent SSc patients, which may have resulted in an over-representation of limited cutaneous SSc. Additionally, the DETECT algorithm was not developed to identify other forms of PH; application of the DETECT algorithm to the total PH population missed 19% of WHO group 2 PH patients, and 37% of WHO group 3 PH patients, both of which are common in SSc. The guideline definitions used for PH classification in DETECT do not consider other variables which may be relevant in clinical practice, for example, PVR, echocardiographic parameters to identify left ventricular disease, exercise haemodynamics or fluid challenge in borderline PH, or elevated transpulmonary gradient that may have increased risk of progression to PAH.[34] Pulmonary veno-occlusive disease was not considered, since neither systematic radiological assessment nor lung biopsy were performed. Finally, the results are based on cross-sectional analyses; it is not possible to determine algorithm performance long-term, or to recommend how frequently patients should be assessed. Results from this study were not validated externally but internal validation using well-established methodology (bootstrapping) confirmed that our findings are robust.

In conclusion, in this cross-sectional multicentre study, we have addressed the fundamental flaw in all previous screening studies in PH by mandating diagnostic RHC in all patients and, thus, determining the false negative rate. The resultant DETECT algorithm is highly sensitive, reducing missed diagnoses when compared with the ESC/ERS guidelines and optimising resource usage by restricting detection efforts to the appropriate high-risk population. Evidence-based guideline recommendations for the identification of mildly symptomatic PAH patients can now be developed, facilitating earlier intervention.

Author affiliations
[1]Cardiology Department, Royal Free Hospital, London, UK
[2]Centre for Rheumatology, Royal Free Hospital, London, UK
[3]Centre for Pulmonary Hypertension, University Hospital, Heidelberg, Germany
[4]Medical University of Vienna, Department of Internal Medicine II, Division of Cardiology, Vienna, Austria
[5]Department of Rheumatology, University Hospital Zurich, Zurich, Switzerland
[6]Department of Internal Medicine, University of Michigan, Ann Arbor, Michigan, USA
[7]Department of Rheumatology and Clinical Immunology, Justus-Liebig-University Giessen, Kerckhoff Clinic Bad Nauheim, Germany
[8]Department of Medicine, Division of Rheumatology, Western University of Canada, London, Ontario, Canada
[9]Department of Rheumatology, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands
[10]Global Medical Affairs, Actelion Pharmaceuticals Ltd, Allschwil, Switzerland
[11]Clinical Development, Actelion Pharmaceuticals Ltd, Allschwil, Switzerland
[12]Center for Medical Statistics, Informatics, and Intelligent Systems, Medical University of Vienna, Vienna, Austria
[13]Scleroderma Research Consultants LLC, Avon, Connecticut, USA

Acknowledgements The authors would like to thank all investigators and patients involved in the DETECT study. They would also like to thank Fabrice Kiefer, PhD and Aisha Rashid, BSc (Actelion Pharmaceuticals Ltd) and Outcome Europe SARL (St-Prex, Switzerland) for their operational support, Juan-Vicente Torres-Martin, MSc (Actelion Pharmaceuticals Ltd) for support of the statistical analyses and the creation of nomograms, Martin Schumacher, PhD (University of Freiburg, Germany) for statistical consultancy, and Julia Heagerty, PhD (Elements Communications Ltd, Westerham, UK) for medical writing support funded by Actelion Pharmaceuticals Ltd.

Contributors All authors except DK were involved in the study design. JGC, CPD, EG, DB, OD, DK, UM-L, JEP and MCV recruited patients to the study and collected data. HC-B and HH conducted the statistical analyses. DR provided epidemiology methods input. All authors had full access to the study data and were involved in the interpretation of results. JGC, JRS and MD drafted the manuscript which was critically reviewed and approved for submission by all authors.

Funding Actelion Pharmaceuticals Ltd funded the study and was responsible for designing the study protocol, data collection, and statistical analysis under the leadership of an independent Study Scientific Committee (non-Actelion authors, apart from HH).

Competing interests JGC has consultancy relationships, received lecture honoraria and/or has received research funding from Actelion Pharmaceuticals Ltd, Pfizer, GlaxoSmithKline and United Therapeutics. CPD has consultancy relationships, received lecture honoraria and/or has received research funding from Actelion Pharmaceuticals Ltd, Pfizer, GlaxoSmithKline, Sanofi-Aventis and Novartis. EG has received honoraria for consultations and/or speaking at conferences from Actelion Pharmaceuticals Ltd, Bayer, Gilead, GlaxoSmithKline, Lilly, Milteney, Novartis, Pfizer and Rotex Medica, and funding for clinical trials from Actelion Pharmaceuticals Ltd, Bayer, GlaxoSmithKline, Encysive, Lilly and Pfizer. DB has acted as a consultant for Actelion Pharmaceuticals Ltd. OD has consultancy relationships and/or has received research funding from Actelion Pharmaceuticals Ltd, Pfizer, Boehringer-Ingelheim, Bayer, Roche, Ergonex, Bristol-Myers Squibb, Sanofi-Aventis, United BioSource Corporation, Medac, Biovitrium, Novartis, 4D Science, Sinoxa and Active Biotec. He has received lecture honoraria from Actelion, Pfizer and Ergonex. DK has consultancy relationships, has served on speakers' bureaus, and/or has received research funding from Actelion Pharmaceuticals Ltd, Bayer, Bristol-Myers Squibb, Gilead, Genentech, ISDIN, Merck, Roche, Sanofi-Aventis and United Therapeutics. UM-L has acted as a consultant and lecturer for Actelion Pharmaceuticals Ltd, Pfizer and GlaxoSmithKline. JEP has received payment for consultancy and lectures and received research funding from Amgen, Bristol-Myers Squibb and Pfizer. She has received payment for consultancy and lectures from Abbott, Janssen, UCB and Roche. She has received consultancy fees and research funding from Actelion Pharmaceuticals Ltd and research funding from Teva and Celgene. MCV has consultancy relationships, received lecture honoraria and/or has received research funding from Actelion Pharmaceuticals Ltd, Pfizer, GlaxoSmithKline, Therabel Pharma and United Therapeutics. MD, HC-B and DMR are full-time employees of Actelion Pharmaceuticals Ltd and have stock/stock options in the company. HH has acted as a consultant for Actelion Pharmaceuticals Ltd and has received research funding from Roche Austria. VVM has acted as a consultant and/or received honoraria/lecture fees from Actelion Pharmaceuticals Ltd, Bayer, Gilead and United Therapeutics. She has received research funding (to the University of Michigan) from Actelion Pharmaceuticals Ltd, Bayer, Novartis and United Therapeutics. JRS has consultancy relationships regarding development of therapies for cardiopulmonary complications of scleroderma with Actelion Pharmaceuticals Ltd, United Therapeutics, Pfizer, Gilead, Bayer, Boehringer-Ingelheim, Sigma Tau, FibroGen, Sanofi, Celgene, MedImmune, Genentech and Intermune. He has received research funding from Actelion Pharmaceuticals Ltd, United Therapeutics and Gilead Sciences and payment for speakers' bureaus from United Therapeutics.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1 Galiè N, Hoeper MM, Humbert M, *et al*. Guidelines for the diagnosis and treatment of pulmonary hypertension: the Task Force for the Diagnosis and Treatment of Pulmonary Hypertension of the European Society of Cardiology (ESC) and the European Respiratory Society (ERS), endorsed by the International Society of Heart and Lung Transplantation (ISHLT). *Eur Heart J* 2009;30:2493–537.
2 McLaughlin VV, Archer SL, Badesch DB, *et al*. ACCF/AHA 2009 expert consensus document on pulmonary hypertension: a report of the American College of Cardiology Foundation Task Force on Expert Consensus Documents and the American Heart Association: developed in collaboration with the American College of Chest Physicians, American Thoracic Society, Inc., and the Pulmonary Hypertension Association. *Circulation* 2009;119:2250–94.
3 Steen VD, Medsger TA. Changes in causes of death in systemic sclerosis, 1972–2002. *Ann Rheum Dis* 2007;66:940–4.
4 Tyndall AJ, Bannet B, Vonk M, *et al*. Causes and risk factors for death in systemic sclerosis: a study from the EULAR Scleroderma Trial and Research (EUSTAR) database. *Ann Rheum Dis* 2010;69:1809–15.

5 Hachulla E, Carpentier P, Gressin V, *et al*. Risk factors for death and the 3-year survival of patients with systemic sclerosis: the French ItinérAIR-Sclérodermie study. *Rheumatology (Oxford)* 2009;48:304–8.
6 Condliffe R, Kiely DG, Peacock AJ, *et al*. Connective tissue disease-associated pulmonary arterial hypertension in the modern treatment era. *Am J Respir Crit Care Med* 2009;179:151–7.
7 Launay D, Sitbon O, Hachulla E, *et al*. Survival in systemic sclerosis-associated pulmonary arterial hypertension in the modern management era. *Ann Rheum Dis* 2013;72:1940–6.
8 Humbert M, Yaici A, de Groote P, *et al*. Screening for pulmonary arterial hypertension in patients with systemic sclerosis: clinical characteristics at diagnosis and long-term survival. *Arthritis Rheum* 2011;63:3522–30.
9 Galiè N, Rubin LJ, Hoeper M, *et al*. Treatment of patients with mildly symptomatic pulmonary arterial hypertension with bosentan (EARLY study): a double-blind randomized controlled trial. *Lancet* 2008;371:2093–100.
10 Badesch DB, Champion HC, Gomez Sanchez MA, *et al*. Diagnosis and assessment of pulmonary arterial hypertension. *J Am Coll Cardiol* 2009;54(Suppl 1):55–66.
11 Williams MH, Handler CE, Akram R, *et al*. Role of N-terminal brain natriuretic peptide (N-TproBNP) in scleroderma-associated pulmonary arterial hypertension. *Eur Heart J* 2006;27:1485–94.
12 Steen V, Medsger TA Jr. Predictors of isolated pulmonary hypertension in patients with systemic sclerosis and limited cutaneous involvement. *Arthritis Rheum* 2003; 48:516–22.
13 Schreiber BE, Valerio CJ, Keir GJ, *et al*. Improving the detection of pulmonary hypertension in systemic sclerosis using pulmonary function tests. *Arthritis Rheum* 2011;63:3531–9.
14 Allanore Y, Borderie D, Avouac J, *et al*. High N-terminal pro-brain natriuretic peptide levels and low diffusing capacity for carbon monoxide as independent predictors of the occurrence of precapillary pulmonary arterial hypertension in patients with systemic sclerosis. *Arthritis Rheum* 2008;58:284–91.
15 Fisher MR, Forfia PR, Chamera E, *et al*. Accuracy of Doppler echocardiography in the hemodynamic assessment of pulmonary hypertension. *Am J Respir Crit Care Med* 2009;179:615–21.
16 Parent F, Bachir D, Inamo J, *et al*. A hemodynamic study of pulmonary hypertension in sickle cell disease. *N Engl J Med* 2011;365:44–53.
17 Wilson JMG, Jungner G. *Principles and practice of screening for disease. Public Health Paper, Number 34*. Geneva: WHO, 1968.
18 Subcommittee for Scleroderma Criteria of the American Rheumatism Association Diagnostic and Therapeutic Criteria Committee. Preliminary criteria for the classification of systemic sclerosis (scleroderma). *Arthritis Rheum* 1980;23: 581–90.
19 Simonneau G, Robbins IM, Beghetti M, *et al*. Updated clinical classification of pulmonary hypertension. *J Am Coll Cardiol* 2009;54(Suppl 1):43–54.
20 Harrell FE Jr, Margolis PA, Gove S, *et al*. Development of a clinical prediction model for an ordinal outcome: the World Health Organization Multicentre Study of Clinical Signs and Etiological agents of Pneumonia, Sepsis and Meningitis in Young Infants. WHO/ARI Young Infant Multicentre Study Group. *Stat Med* 1998;17:909–44.
21 Hachulla E, Gressin V, Guillevin L, *et al*. Early detection of pulmonary arterial hypertension in systemic sclerosis: a French nationwide prospective multicenter study. *Arthritis Rheum* 2005;52:3792–800.
22 Mukerjee D, St George D, Coleiro B, *et al*. Prevalence and outcome in systemic sclerosis associated pulmonary arterial hypertension: application of a registry approach. *Ann Rheum Dis* 2003;62:1088–93.
23 Badesch DB, Raskob GE, Elliott CG, *et al*. Pulmonary arterial hypertension: baseline characteristics from the REVEAL Registry. *Chest* 2010;137:376–87.
24 Phung S, Strange G, Chung LP, *et al*. Prevalence of pulmonary arterial hypertension in an Australian scleroderma population: screening allows for earlier diagnosis. *Intern Med J* 2009;39:682–91.
25 Nihtyanova SI, Tang EC, Coghlan JG, *et al*. Improved survival in systemic sclerosis is associated with better ascertainment of internal organ disease: a retrospective cohort study. *QJM* 2010;103:109–15.
26 Hoeper MM, Lee SH, Voswinckel R, *et al*. Complications of right heart catheterization procedures in patients with pulmonary hypertension in experienced centers. *J Am Coll Cardiol* 2006;48:2546–52.
27 Johnson SR, Fransen J, Khanna D, *et al*. Validation of potential classification criteria for systemic sclerosis. *Arthritis Care Res (Hoboken)* 2012;64:358–67.
28 Thenappan T, Shah SJ, Rich S, *et al*. A USA-based registry for pulmonary arterial hypertension: 1982–2006. *Eur Respir J* 2007;30:1103–10.
29 Cavagna L, Caporali R, Klersy C, *et al*. Comparison of brain natriuretic peptide (BNP) and NT-proBNP in screening for pulmonary arterial hypertension in patients with systemic sclerosis. *J Rheumatol* 2010;37:2064–70.
30 Bossone E, Paciocco G, Iarussi D, *et al*. The prognostic role of the ECG in primary pulmonary hypertension. *Chest* 2002;121:513–18.
31 Jiang X, Han ZY, Wang Y, *et al*. Hemodynamic variables and clinical features correlated with serum uric acid in patients with pulmonary arterial hypertension. *Chin Med J (Engl)* 2008;121:2497–503.
32 Njaman W, Iesaki T, Iwama Y, *et al*. Serum uric acid as a prognostic predictor in pulmonary arterial hypertension with connective tissue disease. *Int Heart J* 2007;48:523–32.
33 Avouac J, Huscher D, Furst DE, *et al*. Expert consensus for performing right heart catheterisation for suspected pulmonary arterial hypertension in systemic sclerosis: a Delphi consensus study with cluster analysis. *Ann Rheum Dis* 2013. Published Online First: 24 Jan 2013. doi:10.1136/annrheumdis-2012-202567
34 Valerio CJ, Schreiber BE, Handler CE, *et al*. Borderline mean pulmonary artery pressure in systemic sclerosis patients: Trans-pulmonary gradient predicts risk of developing pulmonary hypertension. *Arthritis Rheum* 2012;65:1074–84.

Evidence-based detection of pulmonary arterial hypertension in systemic sclerosis: The DETECT study

Web Only file

APPENDICES

	Page
List of investigators	3
Methodology	4
Appendix 1: Analysis sets and patient groups	4
Appendix 2: Complete list of variables (number of variables)	5
Appendix 3: Analytical methods	7
Appendix 4: Variable reduction process by univariable and multivariable analysis	8
Appendix 5: Construction of PAH detection algorithm as a two-step decision tree	13
Appendix 6: Construction of prediction nomograms	20
Appendix 7: Bootstrap validation of models	20
Appendix 8: Sensitivity analyses – handling of missing and sparse data	25
Appendix 9: Sensitivity analyses – influence of omitting a variable in the decision tree	26
Appendix 10: Interpretation of ESC/ERS guidelines	29
Appendix 11: Statistical software	29
References	30

LIST OF INVESTIGATORS

We are indebted to all the local investigators and staff members for their collaboration and commitment, in particular to the following:

Austria: D Bonderman, H Olschewski, G Klein. **Bosnia**: S Sokolovic. **Canada**: JE Pope, P Lee, J-R Goulet, M Baron, E Sutton, J Markland, J Dunne, N Hirani, D Robinson, D Lien, D Langleben. **People's Republic of China**: M Li. **Czech Republic**: R Becvar. **Germany**: E Grünig, U Müller-Ladner, M Aringer, F Behrens, J Distler, P Willeke, C Baerwald, S Rosenkranz, J Thoden, G Riemekasten, D Skowasch, R Ewert, SM Weiner. **Hungary**: I Czuriga/G Szucs, R Faludi/L Czirjak. **Netherlands**: MC Vonk. **Norway**: JT Gran, E Rødevand, L-T Bertelsen. **Poland**: S Sierakowski. **Romania**: S Rednic. **Russia**: E Nasonov. **Slovakia**: J Lukac. **Spain**: M Egurbide, I Castellvi. **Switzerland**: O Distler, D Weilenmann, C Chizzolini. **Turkey**: M Inanc, I Ertenli, Y Kabasakal, S Ozbek. **United Kingdom**: CP Denton. **United States**: K Phillips, N Sandorfi, B Kahaleh, R Simms, A Goldberg, L Shapiro, N Rothfield, J Condemi, S Volkov, RW Martin, G Karpouzas, D Khanna, H Kenney, A Gerbino, T Osborn, J Molitor, ME Csuka, K Highland, M Cadeiras, M Harris, V Steen, L Hummers, G Mitri, G Rios, F Wigley.

METHODOLOGY

Appendix 1: Analysis sets and patient groups

Analysis sets and patient groups were defined as shown in Table S1. Patients were classified as non-pulmonary hypertension (PH), or World Health Organization (WHO) group 1 PH (PAH), WHO group 2 PH (PH due to left heart disease) or WHO group 3 PH (PH due to lung disease/hypoxia), according to current guidelines.[1,2] The WHO group 3 definition was based on Study Scientific Committee consensus.

Table S1. Analysis sets and patient groups

Analysis sets	**Definition**
All enrolled set (N=488)	All patients who had signed the consent form, met eligibility criteria and were enrolled in the study
RHC analysis set (N=466)	All patients with RHC test results who were in the PH group or in the non-PH group, as defined below
PAH analysis set (N=408)	All patients with RHC test results who were in the PAH group or in the non-PH group, as defined below
Patient groups	**Definition**
PH	mPAP ≥25 mm Hg at rest
WHO group 1 PH **(PAH)**	mPAP ≥25 mm Hg at rest and PCWP ≤15 mm Hg
WHO group 2 PH **(PH due to left heart disease)**	mPAP ≥25 mm Hg at rest and PCWP >15 mm Hg
WHO group 3 PH **(PH due to lung disease/hypoxia)**	mPAP ≥25 mm Hg at rest and PCWP ≤15 mm Hg and (FVC <60% or [FVC 60–70% and 'moderate-severe' parenchymal lung disease on HRCT or HRCT not available])
Non-PH	mPAP <25 mm Hg at rest

Abbreviations: FVC, forced vital capacity; HRCT, high resolution computed tomography; mPAP, mean pulmonary arterial pressure; PAH, pulmonary arterial hypertension; PH, pulmonary hypertension; PCWP, pulmonary capillary wedge pressure; RHC, right heart catheterisation; WHO, World Health Organization.

The analyses presented herein pertain to the PAH analysis set using the PAH and the non-PH patient groups only.

Appendix 2: Complete list of variables (number of variables)

Demographic and clinical parameters (68)

Demographics (3)

Gender; Age; Race.

Physical exam (8)

Systolic blood pressure; Diastolic blood pressure; Presence of peripheral oedema; Presence of crackles, rales or wheezing; Body mass index; WHO functional class (I/II vs. III/IV), 6-minute walk distance (6MWD), Borg dyspnoea index.

Systemic sclerosis (SSc) clinical characteristics (12)

SSc disease duration; SSc subtype (diffuse cutaneous vs. limited cutaneous); SSc subtype (overlap/mixed vs. limited cutaneous); Presence of digital ulcers in past 12 months; Modified Rodnan skin score; Physician global assessment scale; Physician skin disease activity assessment last year; Physician skin disease activity assessment last month; Physician skin disease activity assessment current overall; Overall scleroderma compared to 1 month ago; Overall scleroderma compared to 1 year ago; Raynaud's condition score.

Other SSc clinical characteristics (current/past) and medical history (33)

Raynaud's phenomenon; Calcinosis; Telangiectasias; Anginal pain; Syncope/near syncope; Palpitations; Pericarditis; Myocarditis; Dyspnoea; Cough; Proteinuria; Gastro-oesophageal reflux; Dysphagia; Aspiration; Diarrhoea; Biliary cirrhosis; Dyspareunia; Erectile dysfunction;

Arthralgia; Myalgia; Muscle weakness; Tendon friction rub; Loss in joint range of motion; Depression; Sjogren/Sicca syndrome; Fatigue/malaise; Systemic lupus erythematosus; Rheumatoid arthritis; Polymyositis; Dermatomyositis; Mixed connective tissue disease; History of any connective tissue disease; Current/previous smoker.

Pulmonary function tests and haemoglobin (12)

Forced vital capacity (FVC); FVC % predicted; Forced expiratory volume in 1 second (FEV1); FEV1 % predicted; Pulmonary diffusing capacity for carbon monoxide (DLCO); DLCO % predicted; DLCO/alveolar volume (DLCO/VA); DLCO/VA % predicted; Total lung capacity % predicted; Residual volume % predicted; FVC % predicted/DLCO % predicted; Haemoglobin.

Serum laboratory (13)

N-terminal pro-brain natriuretic peptide (NTproBNP); Endothelin-1; von Willebrand factor; C-reactive protein; Creatinine; Estimated glomerular filtration rate; Serum urate; Erythrocyte sedimentation rate; Anti-centromere antibody (ACA); Anti-topoisomerase-I (Scl-70) antibody; Anti-U3-RNP (fibrillarin) antibody; Anti-U1-RNP antibody; Anti-RNA polymerase antibody.

Electrocardiography (3)

Right ventricular strain; Right axis deviation (RAD); Right bundle branch block.

Echocardiography (28)

Aortic root; Left atrium; Inferior vena cava; Interventricular septum; Tricuspid annular plane systolic excursion (TAPSE); Posterior wall; Right atrium (RA) area; Right ventricle (RV) area; RV diameter; Left ventricle (LV) end-diastolic dimension; LV end-systolic dimension; Tissue Doppler imaging (TDI) tricuspid annulus S; TDI tricuspid annulus E'; TDI tricuspid annulus A'; TDI mitral annulus S; TDI mitral annulus E'; TDI mitral annulus A'; Pulsed-wave Doppler

mitral inflow E'; Pulsed-wave Doppler mitral inflow A'; Tricuspid regurgitant jet (TR) velocity; Pulmonary regurgitant velocity; Aortic valve (normal/abnormal); Mitral valve (normal/abnormal); Tricuspid valve (normal/abnormal); Pulmonary valve (normal/abnormal); Pericardial effusion (yes/no); Qualitative assessment of RV pump function; Qualitative assessment of LV pump function.

The total number of variables was 112.

Appendix 3: Analytical methods

Univariable logistic regression (ULR) and multivariable logistic regression (MLR) modelling were the main analytical methods, including linear and non-linear functional relations, where the binary outcome variable was PAH versus non-PH.

Variable selection in MLR was performed in different stages by means of stepwise forward procedure and clinical judgement (see Appendix 4 and Appendix 5).

Discriminatory performance to distinguish between PAH and non-PH patients was examined by receiver operating characteristic (ROC) curve analysis. A ROC curve shows the relationship between the true-positive rate (sensitivity on y-axis) and false-positive rate (1−specificity on x-axis). The ROC area under the curve (AUC), also called the concordance statistic (C-statistic), formed the criterion for assessing the discriminatory ability of a model. A risk prediction model with perfect discrimination (AUC=100%) has a ROC curve that passes through the upper left corner (100% sensitivity, 100% specificity) and pure chance discrimination (AUC=50%) has a ROC curve that is a diagonal line. The ROC AUC and its 95% confidence intervals were calculated for each ULR and MLR model.

A risk cut-off can be calculated to classify subjects as having PAH or non-PH for a fitted model. Among others, this can be done by pre-specifying either sensitivity or specificity

levels. A two-by-two classification table (Table S2) was created for each assessed cut-off showing frequencies, together with the discriminatory performance statistics: sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV), and their 95% confidence intervals, when relevant.

Table S2. Two-by-two classification of diagnostic test characteristics

	PAH (+ve) (by RHC)	Non-PH (-ve) (by RHC)	Total	
Algorithm (+ve) ≥x score cut-off	a (true positive)	b (false positive)	a+b	a/(a+b) **Positive predictive value (PPV)**
Algorithm (–ve) <x score cut-off	c (false negative)	d (true negative)	c+d	d/(c+d) **Negative predictive value (NPV)**
Total	a+c	b+d	a+b+c+d	
	a/(a+c) **Sensitivity**	d/(b+d) **Specificity**		

Note: (a + b)/(a + b + c + d) = RHC referral rate; c/(a + c) = missed PAH diagnoses rate.

Calibration of MLR models was assessed by the Hosmer-Lemeshow χ^2 goodness-of-fit test (HL-test), where a significant p-value implies lack of model fit.

Appendix 4: Variable reduction process by univariable and multivariable analysis

a) Methods

In order to manage a large number of variables (112) relative to the number of PAH patients, four groups of clinically related variables were formed for the model building process:

standard demographic and clinical variables; serum tests; electrocardiography (ECG); and echocardiography (ECHO). Selection of a final set of variables for detection purposes was performed in a series of model building stages to obtain a reduced set of good-performing variables. The first variable selection stage of the large set of potential candidate variables started after the ULR analysis, where only the statistically good-performing variables (Wald χ^2 test statistic p-value <0.15 for linear terms and <0.05 for quadratic terms) were selected and carried forward to the MLR analysis stage. The second variable selection stage consisted of MLR within the groups of variables to further reduce the set of potential candidate variables by stepwise forward selection procedure (SELECTION=STEPWISE method in SAS Proc Logistic; with entry criterion slentry =0.15 and retention criterion slstay =0.10). Interaction effect between two variables was tested for its statistical significance (Wald χ^2 test statistic p-value <0.15), whenever it was suspected to exist by the Study Scientific Committee. A review of the results of statistically good-performing variables was performed by the Study Scientific Committee and variables were further selected based on clinical plausibility and/or feasibility with particular regard to resource limitations in standard real-world practice. Clinical feasibility means that for variables related to the same clinical condition, if they had a similar performance, the easiest to measure in clinical practice was selected.

b) Results

Using the described sequence of statistical analyses, combined with clinical judgment, the initial set of 112 variables was reduced to 13 good-performing and clinically well-accepted variables (Table S3 and Table S4). These thirteen variables formed the basis for constructing the final detection algorithm as a two-step decision tree (see Appendix 5).

Table S3. Variable selection process

Variable selection step	Variable selection criteria	Number of variables selected				Total number of variables selected
		Demographic and clinical	Serum laboratory	ECG variables	ECHO variables	
All variables	–	68	13	3	28	112
Univariable logistic regression	$p<0.15$	25	8	2	12	47
Multivariable logistic regression	Entry criterion $p<0.15$, retention criterion $p<0.10$	7	4	2	4	17
Study Scientific Committee consensus	Clinical plausibility and practical feasibility	4	4	1	4	13
2-step decision tree (multivariable logistic regression)	Entry criterion $p<0.15$, retention criterion $p<0.10$	2	3	1	2	8

Table S4. Univariable analysis of the 13 selected variables of which eight were used for the DETECT algorithm

Variables	Summary Statistics		Univariable Logistic Regression	
	Non-PH patients	PAH patients	Wald Chi-square (p-value)	ROC AUC, % (95% CI)
Demographic and clinical characteristics				
Systemic sclerosis characteristics Current/past telangiectasias, n/N (%)	218/321 (67.9)	76/87 (87.4)	11.82 (<0.001)	59.7 (55.4, 64.1)
Physical examination Peripheral oedema present, n/N (%)	39/318 (12.3)	19/87 (21.8)	4.97 (0.026)	54.8 (50.0, 59.5)
WHO functional class, n/N (%) I/II III/IV	256/306 (83.7) 50/306 (16.3)	56/87 (64.4) 31/87 (35.6)	14.66 (<0.001)	59.6 (54.2, 65.1)
Pulmonary function tests, n FVC % predicted/DLCO % predicted mean (SD) median (Q1, Q3)	321 1.8 (0.5) 1.7 (1.4, 2.0)	87 2.2 (0.7) 2.1 (1.7, 2.5)	33.19 (<0.001)	71.5 (65.6, 77.4)
Serum tests				
n NTproBNP, log_{10} mean (SD) median (Q1, Q3) Serum urate, mg/100 mL mean (SD) median (Q1, Q3)	306 2.1 (0.5) 2.1 (1.8, 2.3) 4.7 (1.5) 4.5 (3.7, 5.4)	80 2.4 (0.5) 2.3 (2.0, 2.8) 5.9 (1.5) 5.8 (4.7, 6.8)	25.42 (<0.001) 30.16 (<0.001)	67.5 (60.9, 74.2) 71.9 (65.9, 77.9)

Anti-centromere antibody positive, n/N (%) Scl-70 positive, n/N (%)	77/306 (25.2) 91/299 (30.4)	40/80 (50.0) 12/79 (15.2)	17.64 (<0.001) 6.99 (0.008)	62.4 (56.4, 68.4) 57.6 (52.9, 62.4)
Electrocardiography				
Right axis deviation* present, n/N (%)	10/291 (3.4)	11/83 (13.3%)	10.19 (0.001)	54.9 (51.1, 58.7)
Echocardiography				
TAPSE, mm, n mean (SD) median (Q1, Q3)	284 22.7 (4.8) 23.0 (20.0, 26.0)	76 21.5 (4.3) 21.0 (19.0, 24.0)	3.92 (0.048)	59.0 (51.7, 66.2)
Right atrium area, cm^2, n mean (SD) median (Q1, Q3)	286 13.4 (4.7) 12.6 (10.2, 15.4)	82 17.1 (6.2) 15.7 (13.9, 19.0)	24.45 (<0.001)	71.2 (65.0, 77.3)
Right ventricle area, cm^2, n mean (SD) median (Q1, Q3)	291 15.0 (5.4) 14.7 (12.0, 17.4)	82 19.3 (6.8) 18.4 (13.7, 22.6)	25.81 (<0.001)	69.3 (62.5, 76.0)
TR velocity, m/s, n mean (SD) median (Q1, Q3)	303 2.4 (0.6) 2.4 (2.1, 2.7)	84 3.0 (0.8) 3.0 (2.5, 3.5)	62.28 (<0.001)	79.5 (73.7, 85.3)

*QRS axis ≥90°. Abbreviations: AUC, area under the curve; DLCO, pulmonary diffusing capacity for carbon monoxide; FVC, forced vital capacity; NTproBNP, N-terminal pro-brain natriuretic peptide; PAH, pulmonary arterial hypertension; PH, pulmonary hypertension; ROC; receiver operating characteristic; TAPSE, tricuspid annular plane systolic excursion; TR, tricuspid regurgitant jet; WHO, World Health Organization.

Appendix 5: Construction of PAH detection algorithm as a two-step decision tree

a) Methods

The 13 variables selected as good-performers during the above stages of model building were divided into nine non-ECHO variables and four ECHO variables to align the detection algorithm with real-world practice where the rheumatologist managing the patient accesses non-ECHO data prior to referral to a cardiologist for echocardiography. These two groups of variables formed the basis for developing a two-step detection algorithm as a decision tree, as shown in Figure 2 of the manuscript.

The development of the decision tree entailed building two models by MLR; the nine non-ECHO variables were candidate variables in the first step of the decision tree and the four ECHO variables were candidate variables in the second step of the decision tree. In order to carry forward non-ECHO PAH risk information, the step 1 linear risk prediction score was included at step 2 (Figure S1). The decision tree allows patients with a low risk of PAH to be filtered out in the first step, which saves further assessments by the cardiologist. A stepwise forward selection procedure was applied within each step where the variable selection criteria were: entry criterion 0.15 and retention criterion 0.10.

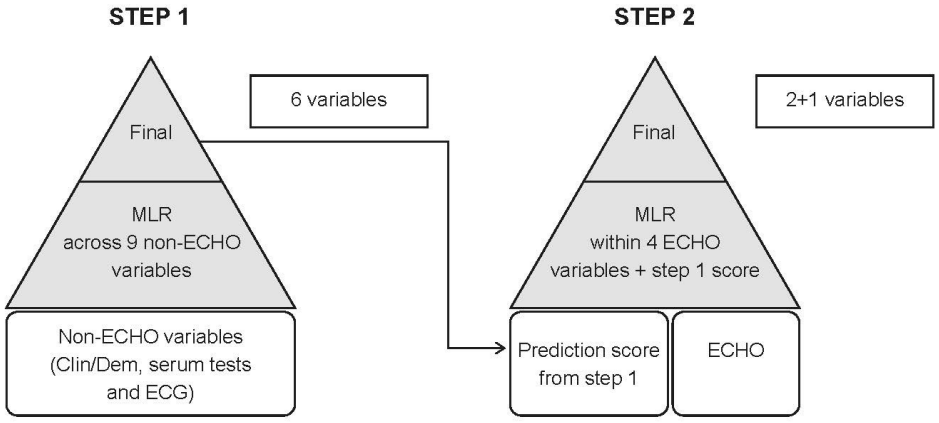
No significant lack of fit was observed for the final MLR model at step 1 (HL p-value =0.23), nor for the final MLR model at step 2 (HL p-value =0.74).

Quadratic terms for the continuous variables were included in the variable selection process in a hierarchical manner, that is, selection of a quadratic term implies selection of the corresponding linear term as well. For those variables where the quadratic term was significant, different approaches were applied to further examine non-linear relationships using: flexible non-parametric methods (Generalized Additive Models [GAM] with penalised

splines,[3] Restricted Cubic Splines,[4] or variable transformations [e.g., $\log_{10}$ or quintiles categorisation]).

Based on the prediction scores of the two MLR models of the decision tree, cut-offs were selected to classify patients for referral to echocardiography at step 1, and for referral to RHC at step 2. As there is a trade-off between sensitivity and specificity, different levels of sensitivity and specificity can be selected to identify cut-offs to achieve acceptable global performance characteristics of the detection algorithm. The objective of step 1 was to achieve a low rate of missed PAH diagnoses (i.e. low false-negative rate) by selecting a high sensitivity, which was fixed at 97%. In step 2, the level of specificity was selected in order to obtain a high PPV (high rate of positive PAH diagnoses). Therefore, a range of specificities (35% to 85%) was used to examine the impact on the rate of missed diagnoses.

Figure S1. Construction of the two-step decision tree algorithm



Abbreviations: Clin/Dem, demographic and clinical parameters; ECG, electrocardiography; ECHO, echocardiography; MLR, multivariable logistic regression.

b) Results

The forward stepwise selection procedure reduced the number of variables from nine to six in step 1 and from four to two in step 2. None of the statistically selected variables was replaced by another non-ECHO variable by the Study Scientific Committee at step 1, whereas one ECHO variable (RV area) was replaced by another ECHO variable (RA area) as the latter is regarded as easier to measure and likely to be more reproducible. This replacement had a minimal effect on the performance of the step 2 model (AUC of 89% and 88% using RV area and RA area, respectively).

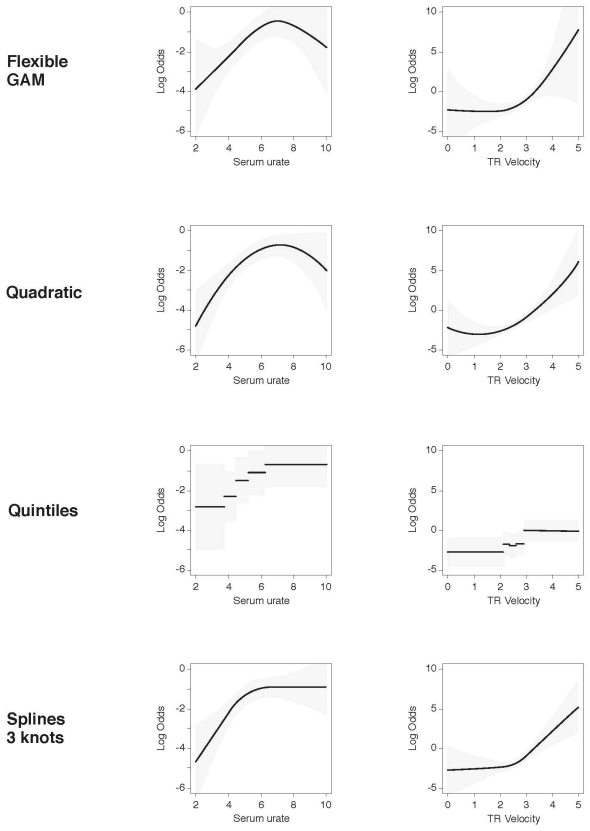
For serum urate at step 1 and TR velocity at step 2, it was necessary to fit a non-linear relationship. Different approaches were applied to adjust for non-linearity: flexible GAM, quadratic function, categorisation with quintiles, and flexible spline functions (Figure S2). Among all these options, no relevant differences were observed in terms of discrimination (same AUC=84% for all approaches). Splines with three knots were finally used in order to avoid the inversion of PAH risk at the extremes of the curves, as recommended by the Study Scientific Committee.

MLR model coefficient estimates and their statistical significance (Wald test) are given in Table 2 of the main manuscript. A positive relationship was observed between all final variables and PAH risk, as indicated by the positive value of the regression coefficients of the main terms.

The global performance of the detection algorithm at a pre-defined sensitivity of 97% at step 1 and different specificity levels at step 2 is given in Table S5. This table shows that as the specificity at step 2 increases, the global specificity and global PPV also increase. The table also shows the natural trade-off between sensitivity and specificity, i.e., as one increases the other decreases.

Characteristics of patients that were misdiagnosed as non-PAH by the decision algorithm are presented in Table S6. At step 1 (97% sensitivity) two out of 77 PAH patients who had complete data for step 1 were misdiagnosed. For a range of specificity levels between 35% and 85% at step 2, the higher the specificity level the greater the number of missed diagnoses. Characteristics of the 24 patients with missed PAH diagnoses based on current current ESC/ERS guidelines are presented in Table S7.

Figure S2. Different approaches to assess non-linear functional relationships of serum urate (step 1) and TR velocity (step 2) of the decision tree



Partial fitting for serum urate at step 1 (left column) and TR velocity at step 2 (right column) with different non-linear function. The Y-axis represents the log-odds of having PAH in the MLR model. Abbreviations: GAM, Generalized Additive Models; TR, tricuspid regurgitant jet.

Table S5. Global performance of the final two-step decision tree algorithm at 97% sensitivity at step 1 and different levels of specificity at step 2

Specificity step 2 (%)	Overall sensitivity (%)	Overall specificity (%)	Overall PPV (%)	Overall NPV (%)	Overall missed PAH diagnoses (false negatives, %)
35	96	48	35	98	4
40	89	52	35	94	11
45	89	56	37	94	11
50	88	60	39	94	12
55	88	64	41	95	12
60	88	68	44	95	12
65	85	72	47	94	15
70	83	78	50	94	17
75	80	80	54	93	20
80	80	83	60	94	20
85	74	88	64	92	26

Abbreviations: NPV, negative predictive value; PAH, pulmonary arterial hypertension; PPV, positive predictive value.

Table S6. Details of the patients with missed PAH diagnoses using a sensitivity of 97% at step 1 and a range of specificities at step 2

			Step 1 variables						Step 2 variables	
Step	**Criteria**	**Cumulative number of missed PAH diagnoses**	**RAD (presence)**	**Telang. (presence)**	**ACA (presence)**	**NT-proBNP**	**FVC % pred./DLCO % pred.**	**Serum urate**	**RA area**	**TR velocity**
1	Sens 97%	2	No	Yes	No	204	1.70	2.3	16.4	2.90
	Sens 97%		No	No	No	47	1.74	3.7	15.0	2.52
2	Spec 35%	+1=3	No	Yes	No	71	2.21	3.8	8.5	2.59
	Spec 45%	+5=8	No	Yes	No	25	1.70	5.2	13.0	2.30
			No	Yes	No	47	1.32	4.6	17.7	2.52
			No	Yes	No	81	1.32	3.6	18.0	2.90
			No	No	Yes	210	1.45	6.1	10.0	2.14
			No	Yes	No	102	1.38	4.3	13.0	2.71
	Spec 60%	+1=9	No	Yes	No	44	2.37	4.9	8.8	2.25
	Spec 65%	+2=11	No	Yes	Yes	71	1.46	3.5	7.0	3.20
			No	Yes	Yes	46	2.07	4.3	16.7	2.00
	Spec 70%	+1=12	No	Yes	No	187	2.29	7.3	11.5	1.06
	Spec 80%	+2=14	No	Yes	No	41	1.85	6.3	14.0	2.78
			No	Yes	No	155	2.04	5.4	15.8	2.40
	Spec 85%	+5=19	No	Yes	Yes	95	2.29	4.5	10.0	2.75
			No	Yes	Yes	135	1.86	5.5	16.0	2.25
			No	Yes	Yes	215	1.92	5.2	10.3	2.67
			No	Yes	No	287	1.96	7.6	16.9	2.53
			No	Yes	Yes	166	1.78	6.3	11.4	2.65

Abbreviations: ACA, anti-centromere antibody; DLCO, pulmonary diffusing capacity for carbon monoxide; FVC, forced vital capacity; NTproBNP, N-terminal pro-brain natriuretic peptide; PAH, pulmonary arterial hypertension; RA, right atrium; RAD, right axis deviation; Telang; telangiectasias; TR, tricuspid regurgitant jet.

Table S7. Characteristics of the 24 patients with missed PAH diagnoses using the ESC/ERS guidelines[1]

mPAP (mm Hg)	PCWP (mm Hg)	PVR (dyn·sec/cm^5)	TR velocity (m/s)	RA area (cm^2)	NTproBNP (pg/mL)	RAD (presence)	ACA (presence)
36	12	400.00	3.30	15.5	123	No	Yes
25	12	208.00	2.50	15.8	55	-	No
25	11	169.70	*	14.8	92	Yes	No
26	10	261.22	3.04	14.4	-	No	-
31	5	520.00	3.35	12.0	175	Yes	Yes
26	14	147.69	2.14	10.0	210	No	Yes
27	14	231.11	1.06	11.5	187	No	No
26	10	266.67	*	12.5	160	No	Yes
26	10	196.92	2.78	14.0	41	No	No
30	9	236.62	2.47	13.4	-	-	-
27	13	200.00	*	14.8	191	No	Yes
29	13	272.34	2.62	15.2	133	No	Yes
28	14	169.70	2.65	11.4	166	No	Yes
25	9	237.04	2.75	10.0	95	No	Yes
26	13	226.09	3.00	13.9	383	No	No
30	12	257.14	2.59	8.5	71	No	No
28	13	260.87	2.40	15.8	155	No	No
25	13	160.00	2.90	18.0	81	No	No
25	5	210.53	3.15	15.0	258	No	No
29	8	254.55	2.30	13.0	25	No	No
30	14	312.20	2.71	13.0	102	No	No
25	11	238.30	2.52	15.0	47	No	No
26	15	275.00	2.90	-	-	No	-
37	15	320.00	*	16.0	135	No	Yes

Abbreviations: ACA, anti-centromere antibody; mPAP, mean pulmonary arterial pressure; NTproBNP, N-terminal pro-brain natriuretic peptide; PCWP, pulmonary capillary wedge pressure; PVR, pulmonary vascular resistance; RA, right atrium; RAD, right axis deviation; TR, tricuspid regurgitant jet.

[1] As defined in Appendix 10. * TR not detectable.

Appendix 6: Construction of prediction nomograms

Prediction nomograms were developed for the algorithm in order to provide a tool that can be used in clinical practice (Figure 3 in manuscript).

a) Methods

Model functions of the two-step decision tree were graphically represented as two prediction nomograms. In conventional nomograms, the risk score is left-aligned starting from a patient with the smallest possible value for each covariable. In such a nomogram, highly specific covariables (primarily identifying non-PH) can contribute many points for clinically normal findings, compared with sensitive covariables that primarily identify PAH cases. Although this is methodologically correct, this feature of the conventional nomogram makes it difficult for physicians to interpret the clinical influence of each covariable on the risk of PAH. To solve this problem, we refined the nomogram for step 1 of the algorithm by centring each covariable at the mean value of non-PH patients (except for current/past telangiectasias where the majority of patients had ‘presence’ but we chose to centre around ‘absence’ as this is closer to a healthy patient).

b) Results

See Figure 3 in manuscript.

Appendix 7: Bootstrap validation of models

a) Methods

Internal validation[5,6] was used to assess over-optimism of the fitted models that can occur when a large set of potential candidate predictor variables are fitted with a relatively small

set of cases. Bootstrap re-sampling method[7] was used for internal validation of the model building procedure starting with 12 variables. Bootstrapping was conducted on 12 variables rather than 13 because right ventricle area was eliminated in the final model for reasons of feasibility in clinical practice. Therefore, right ventricle area was not included in the bootstrap validation process.

The bootstrap process was conducted by repeating the following steps 2000 times:

i. Generation of a training sample by random selection with replacement from the original data. A validation sample was generated with all the patients not included in the training sample, approximately one-third of the patients

ii. Variable selection and model fitting was performed as in Appendix 3 using the training sample, where sensitivity was fixed 97% at step 1 and specificity at 35% for step 2

iii. For validation purposes, diagnostic performance statistics (ROC AUC, sensitivity, specificity, PPV and NPV) were calculated for each step of the decision tree and overall using the validation sample.

Mean and 95% confidence intervals for ROC AUC, sensitivity, specificity, PPV and NPV of the 2000 validation samples were computed.

Given that the diagnostic performance statistics were calculated using patients (validation samples) independent from those used for model fitting (training samples), these values can be considered as surrogates for the diagnostic performance of the two-step decision tree applied to new patients outside of the DETECT study.[6]

The proportion of times a potential candidate variable was selected in the 2000 training decision tree models is informative about its robustness as a predictor in the decision tree

(more than 30% indicates a weak, more than 50% a moderately strong, and more than 70% a strong predictor[8,9]).

b) Results

The results of the bootstrap validation process are presented in Table S8 and Table S9. High ROC AUC estimates (~80%), were obtained for both steps of the decision tree algorithm using the validation samples. The overall sensitivity was 87% (95% CI: 71%, 100%) and the overall specificity 53% (95% CI: 35%, 72%; Table S8). All selected variables in the final two-step decision tree algorithm were selected more than 60% of the times in the bootstrap process, indicating high model robustness. The exception was RA area, which was selected 48% of the times (Table S9). The reason for this is that RA area replaced RV area in the final model as a result of a clinical decision by the Study Scientific Committee.

The mean ROC AUC in the bootstrap validation samples (Table S8) at step 1 and step 2 of the decision tree (79% and 83%, respectively) was slightly lower than that in the final models (84% and 88%, respectively). The sensitivity at step 1 was also lower in the validation samples versus the final model (94% vs. 97%). The specificity at step 2 was slightly lower in the validation sample than in the final model (33% vs. 35%). All these measures are slightly lower in the bootstrap validation samples than in the final models, as expected.

Table S8. Summary of diagnostic performance statistics obtained in the bootstrap process for the two-step decision tree using the validation samples

	ROC AUC, % Mean (95% CI)	Sensitivity, % Mean (95% CI)	Specificity, % Mean (95% CI)	PPV, % Mean (95% CI)	NPV, % Mean (95% CI)
Step 1 validation	79 (71, 87)	94 (79, 100)	27 (8, 54)	27 (18, 36)	95 (88, 100)
Step 2 validation	83 (73, 91)	94 (83, 100)	33 (19, 49)	35 (24, 47)	93 (81, 100)
Overall validation	Not available	87 (71, 100)	53 (35, 72)	35 (24, 47)	94 (88, 100)

Abbreviations: AUC, area under the curve; CI, confidence interval; NPV, negative predictive value; PPV, positive predictive value.

The design features of the decision tree imply that the overall sensitivity is smaller than or equal to the intermediate sensitivities of steps 1 and 2, that the overall specificity is larger than or equal to the intermediate specificities of steps 1 and 2, and that the overall PPV equals the intermediate PPV of step 2.

Table S9. Number of times that each variable was selected to be included in the final decision tree algorithm by means of bootstrap using 2000 training samples

	% of times that a variable was selected in the bootstrap process	
Step 1 variables	**Linear term**	**Quadratic term**
FVC % pred./DLCO % pred*	99	52
Serum urate*	97	75
Telangiectasias*	91	–
NTproBNP (log_{10})*	86	25
Right axis deviation*	81	–
Serum ACA*	62	–
Peripheral oedema	28	–
WHO functional class	26	–
Serum Scl-70	16	–
Step 2 variables		
TR velocity*	100	69
Right atrium area*	48	5
TAPSE	24	13

*Variable in the final model. Abbreviations: ACA, anti-centromere antibody; DLCO, pulmonary diffusing capacity for carbon monoxide; FVC, forced vital capacity; NTproBNP, N-terminal pro-brain natriuretic peptide; TR, tricuspid regurgitant jet; TAPSE, tricuspid annular plane systolic excursion; WHO, World Health Organization.

Appendix 8: Sensitivity analyses – handling of missing and sparse data

a) Methods

Analyses were performed on available data (see below TR velocity) and each model fit used all patients who had available data for each of the candidate variables. Because the reporting of some data was not mandatory, some variables (e.g., 6MWD) had a large amount of missing data and were excluded from the multivariable models, thus optimising the use of the maximum number of patients for the clinically important variables in model building.

Categorical variables which had sparse data in some categories (e.g., WHO functional class and SSc sub-type) were re-categorised into clinically meaningful categories to improve model estimation.

For patients in whom TR velocity was reported to be absent and no specific value was provided, the value was imputed as the mean of all available values ≤2.8 m/s within the PAH and within the non-PH groups. This method of imputation was applied in consensus with the Study Scientific Committee. The imputed TR velocity values were used in the main regression analyses: sensitivity analyses without imputation were performed to confirm the suitability of imputed values. The reason for using imputed values was to minimise the loss of patients available for constructing the detection algorithm since TR velocity is a strong predictor.

The effect of the imputation of TR velocity was evaluated comparing the performance of the decision tree algorithm with and without imputation.

b) Results

Since TR velocity is used in step 2 of the decision tree, the effect of using imputation or not was only observed in this step and in the overall performance. The global performance results with and without imputation (97% sensitivity in step 1 and 35% specificity at step 2) are shown in Table S10 below.

Table S10. Global discrimination performance with and without TR velocity imputation (97% sensitivity at step 1 and 35% specificity at step 2)

	Overall sensitivity (%)	Overall specificity (%)	Overall PPV (%)	Overall NPV (%)	Overall missed PAH diagnoses (false negatives, %)
With TR velocity imputation (n=356)	96	48	35	98	4
Without TR velocity imputation (n=336)	95	48	35	97	5

Abbreviations: NPV, negative predictive value; PAH, pulmonary arterial hypertension; PPV, positive predictive value; TR, tricuspid regurgitant velocity.

Appendix 9: Sensitivity analyses – influence of omitting a variable in the decision tree

a) Methods

The performance of the global decision tree algorithm was evaluated leaving out one variable at a time in order to assess the effect of a missing value in real practice. The omitted variable was imputed using the mean value of non-PH patients.

b) Results

The results of using the decision tree algorithm with one missing variable at a time are shown in Table S11. The effect of imputing a missing variable on the overall performance is related to the discriminatory ability of the variable, showing a worse performance when the missing variable has higher discriminatory ability (e.g., TR velocity). Overall the effect was small.

Table S11. Global discrimination performance when leaving out one variable at a time (97% sensitivity at step 1 and 35% specificity at step 2)

	Overall sensitivity (%)	Overall specificity (%)	Overall PPV (%)	Overall NPV (%)	RHC referral rate (%) (positive screening tests/ all patients)	Overall missed PAH diagnoses (false negatives, %)
No missing variables	96	48	35	98	62	4
Excluding one variable*						
Right axis deviation	96	50	36	98	61	4
Current/past telangiectasias	96	38	31	97	70	4
Serum ACA	93	52	36	96	58	7
Serum NTproBNP	94	47	34	97	62	6
FVC % pred./DLCO % pred.	93	56	38	96	55	7
Serum urate	94	39	31	96	69	6
Right atrium area	95	44	32	97	64	5
TR velocity	92	46	33	95	62	8

*The omitted variable was imputed using the mean value of non-PH patients. Abbreviations: ACA, anti-centromere antibody; DLCO, pulmonary diffusing capacity for carbon monoxide; FVC, forced vital capacity; NPV, negative predictive value; NTproBNP, N-terminal pro-brain natriuretic peptide; PAH, pulmonary arterial hypertension; PPV, positive predictive value; RHC, right heart catheterisation; TR, tricuspid regurgitant jet.

Appendix 10: Interpretation of ESC/ERS guidelines

For performance comparisons of the DETECT model and application of the ESC/ERS guidelines to the DETECT population, the ESC/ERS guidelines were interpreted as RHC being recommended if any of the following situations was present:

1. TR velocity >3.4 m/s
2. TR velocity >2.8–≤3.4 m/s **AND** symptomatic (defined as at least one of the following DETECT parameters: current anginal pain, current syncope/near syncope, current dyspnoea, presence of peripheral oedema)
3. TR velocity ≤2.8 m/s **AND** symptomatic (defined as above) **AND** presence of additional echocardiography variables suggestive of PH (right atrium area >16 cm^2 and/or ratio of right ventricular diameter/left ventricular end diastolic diameter >0.8).

Appendix 11: Statistical software

Statistical analyses were performed using SAS® (versions 9.2 and 9.3) for descriptive analyses and variables reduction process. The R-software (version 2.13.2) was used mainly for the decision tree modelling. Logistic regression models were fitted using 'Function lrm' from the R package 'Design'. Nomograms were produced using a modified version of 'Function nomogram' from the Design package, in order to obtain the 'centred' nomogram (Appendix 6).

REFERENCES

1. Badesch DB, Champion HC, Sanchez MA, et al. Diagnosis and assessment of pulmonary arterial hypertension. *J Am Coll Cardiol* 2009; **54:** S55–S66.

2. Simonneau G, Robbins IM, Beghetti M, et al. Updated clinical classification of pulmonary hypertension. *J Am Coll Cardiol* 2009; **54:** S43–S54.

3. Wood SN. Generalized Additive Model. An Introduction with R. Texts in Statistical Science. Boca Raton, Fl: Chapman & Hall/CRC; 2006. ISBN 1-58488-474-6.

4. Harrell FE Jr. Regression modeling strategies. Springer series in statistics. Springer, 2001.

5. Harrell FE Jr, Margolis PA, Gove S, et al. Development of a clinical prediction model for an ordinal outcome: the World Health Organization Multicentre Study of Clinical Signs and Etiological agents of Pneumonia, Sepsis and Meningitis in Young Infants. WHO/ARI Young Infant Multicentre Study Group. *Stat Med* 1998; **17:** 909–44.

6. König IR, Malley JD, Weimar C, Diener HC, Ziegler A; German Stroke Study Collaboration. Practical experiences on the necessity of external validation. *Stat Med* 2007; **26:** 5499–511.

7. Efron B and Tibshirani R. An introduction to the bootstrap. Monographs on statistics and applied probability. Boca Raton, Fl: Chapman & Hall/CRC, 1993.

8. Chen C-H, George SL. The bootstrap and identification of prognostic factors via Cox's proportional hazards regression model. *Stat Med* 1985; **4:** 39–46.

9. Sauerbrei W, Schumacher M. A bootstrap resampling procedure for model building: application to the Cox regression model. *Stat Med* 1992; **11:** 2093–109.