



# Residual risk identified in routine noninvasive follow-up assessments in pulmonary arterial hypertension

Jonna Ostermann <sup>1</sup>, Julian Pott<sup>1</sup>, Jan K. Hennigs <sup>1</sup>, Kevin Roedel <sup>2</sup>, Christoph Sinning<sup>3</sup>, Lars Harbaum <sup>1,4</sup> and Hans Klose<sup>1,4</sup>

<sup>1</sup>Division of Respiratory Medicine and Centre of Pulmonary Arterial Hypertension Hamburg, Department of Medicine II, University Medical Centre Hamburg-Eppendorf, Hamburg, Germany. <sup>2</sup>Department of Intensive Care Medicine, University Medical Centre Hamburg-Eppendorf, Hamburg, Germany. <sup>3</sup>Department of Cardiology, University Heart and Vascular Centre Hamburg, University Medical Centre Hamburg-Eppendorf, Hamburg Germany. <sup>4</sup>These authors contributed equally.

Corresponding author: Lars Harbaum ([l.harbaum@uke.de](mailto:l.harbaum@uke.de))



Shareable abstract (@ERSpublications)

A systematic approach using a supervised feature selection algorithm indicates that residual mortality risk can be captured from noninvasive assessments during routine follow-up in patients with pulmonary arterial hypertension <https://bit.ly/3ZFfoDR>

**Cite this article as:** Ostermann J, Pott J, Hennigs JK, *et al.* Residual risk identified in routine noninvasive follow-up assessments in pulmonary arterial hypertension. *ERJ Open Res* 2023; 9: 00072-2023 [DOI: 10.1183/23120541.00072-2023].

Copyright ©The authors 2023

This version is distributed under the terms of the Creative Commons Attribution Non-Commercial Licence 4.0. For commercial reproduction rights and permissions contact [permissions@ersnet.org](mailto:permissions@ersnet.org)

Received: 5 Feb 2023  
Accepted: 28 March 2023

## Abstract

**Background** The 2022 ESC/ERS guidelines on pulmonary hypertension recommend noninvasive risk assessments based on three clinical variables during follow-up in patients with pulmonary arterial hypertension (PAH). We set out to test whether residual risk can be captured from routinely measured noninvasive clinical variables during follow-up in PAH.

**Methods** We retrospectively studied 298 incident PAH patients from a German pulmonary hypertension centre who underwent routine noninvasive follow-up assessments including exercise testing, echocardiography, electrocardiography, pulmonary function testing and biochemistry. To select variables, we used least absolute shrinkage and selection operator (LASSO)-regularised Cox regression models. Outcome was defined as mortality or lung transplant after first follow-up assessment.

**Results** 12 noninvasive variables that were associated with outcomes in a training sub-cohort (n=208) after correction for multiple testing entered LASSO modelling. A model combining seven variables discriminated 1-year (area under the curve (AUC) 0.83, 95% confidence interval (CI) 0.68–0.99,  $p=8.4\times 10^{-6}$ ) and 3-year (AUC 0.81, 95% CI 0.70–0.92,  $p=2.9\times 10^{-8}$ ) outcome status in a replication sub-cohort (n=90). The model's discriminatory ability was comparable to that of the guideline approach in the replication sub-cohort. From the individual model components, World Health Organization functional class, 6-min walking distance and the tricuspid annular plane systolic excursion to systolic pulmonary arterial pressure (TAPSE/sPAP) ratio were sensitive to treatment initiation. Addition of TAPSE/sPAP ratio to the guideline approach numerically increased its ability to discriminate outcome status.

**Conclusion** Our real-world data suggest that residual risk can be captured by noninvasive clinical procedures during routine follow-up assessments in patients with PAH and highlights the potential use of echocardiographic imaging to refine risk assessment.

## Introduction

Pulmonary arterial hypertension (PAH) is a rare pulmonary vasculopathy defined by increased pulmonary arterial pressure and vascular resistance [1]. The complex pathology involves progressive and obliterating remodelling of small pulmonary arteries [2]. The most important consequence of PAH is right-sided heart failure through chronically increased afterload to the right ventricle (RV), which causes a complex clinical syndrome affecting multiple organ systems through low cardiac output and systemic venous congestion, and results in high mortality if left untreated [3–5]. Clinical markers that relate not only to pulmonary vascular disease but also to the systemic consequences of PAH were able to predict disease severity and outcomes, and to inform treatment strategies with targeted drugs [1, 6–8].



Multi-variable risk models combining variables from noninvasive and mixed invasive/noninvasive clinical procedures can be used to stratify patients into risk groups for short-term mortality before and during treatment [6–10]. The 2015 and 2022 European Society of Cardiology (ESC)/European Respiratory Society (ERS) guidelines on pulmonary hypertension (PH) recommend risk assessment at diagnosis of PAH using a three-strata model that incorporates up to 18 different variables from clinical observations and mixed invasive/noninvasive procedures to classify patients at low, intermediate and high 1-year mortality risk [1, 11]. A limitation of this three-strata risk-assessment model is that up to 70% of the patients are classified at intermediate risk [1, 8, 10]. A four-strata model based on an abbreviated list of noninvasive clinical measures (Comparative, Prospective Registry of Newly Initiated Therapies for Pulmonary Hypertension (COMPERA) 2.0) was subsequently developed that provided better discrimination within the intermediate-risk group [12, 13]. This model was implemented in the recent version of the 2022 ESC/ERS guidelines and recommended to be used during follow-up visits in patients with PAH [1, 12, 13]. The model included two markers of exercise capacity (World Health Organisation functional class (WHO-FC) and 6-min walking distance (6MWD)) and one biochemical marker relating to myocardial stress (brain natriuretic peptide/N-terminal pro-brain-type natriuretic peptide (BNP/NT-proBNP)). Models based on these three measures, however, achieved only fair to good discriminatory accuracy to distinguish survivors from non-survivors (C-indices <0.80) [6–8, 10, 12–16].

With increasing therapeutic options that are becoming available to patients with PAH, clinical decisions should be guided by the best possible estimates for future outcome events and by risk assessments with sufficient granularity. This includes decisions on far-reaching therapeutic options such as the need for parenteral prostanoid therapy and evaluation for lung transplantation. Particularly, active listing for lung transplantation requires precise estimates for short-term mortality with and without transplantation [17]. A complete knowledge of all possible contributors to mortality risk including cut-off values holds the promise to enhance the predictive ability of contemporary risk stratification tools. Candidate-driven approaches have shown that the use of additional variables derived from echocardiography, cardiopulmonary exercise testing or gas exchange improved risk prediction [9, 18–20]. In addition, advances in machine learning methods might provide a rigorous framework to further refine risk models [21].

We set out to systematically assess the ability of clinical measures obtained from noninvasive procedures in routine follow-up visits to predict outcomes in patients with PAH. Using a supervised feature selection algorithm, we fitted a combinatory model that best predicted 1-year prognosis. Sensitivity to aetiology of PAH and to changes from baseline to follow-up were evaluated, and each model component was assessed individually.

## Methods

### *Selection of patients*

We retrospectively accessed data from all patients with newly diagnosed idiopathic, heritable, drug- or toxin-induced and associated PAH (connective tissue disease (CTD), infection with human immunodeficiency viruses (HIV), porto-pulmonary hypertension, congenital heart disease (CHD)) attending the University Medical Centre Hamburg-Eppendorf, Hamburg, Germany, between 1 January 2009 and 31 December 2019. PAH was diagnosed based on right heart catheterisation with a mean pulmonary arterial pressure (mPAP)  $\geq 25$  mmHg and pulmonary arterial wedge pressure (PAWP)  $\leq 15$  mmHg following the exclusion of other causes of precapillary PH. Diagnosis and classification of disease were consensus-based in an interdisciplinary team. Patients were commenced on an (at the time) approved PAH-targeted drug(s), and follow-up visits were scheduled every 3–6 months (or whenever clinically indicated). The combined primary end-point was all-cause mortality or lung transplant. Follow-up was censored on 31 December 2021. Data collection was approved by the local ethics committee and complied with the Declaration of Helsinki.

### *Data collection*

Clinical data were collected at diagnosis (within 3 months from diagnostic right heart catheterisation) and at the first comprehensive follow-up visit. Age at diagnosis, self-reported sex, height, weight, comorbidities and WHO-FC were recorded along with data from six clinical procedures (right heart catheterisation, echocardiography, electrocardiogram, pulmonary function test, 6-min walking test and routine biochemistry). A detailed list of all pre-defined noninvasive data points, rate of missing data and formulas for secondary data are provided in supplementary table S1. On average, 7.9% of data were missing at follow-up assessments. Data points from available clinical procedures were curated by two independent attending physicians and excluded if deemed implausible by consensus. If biochemical measurements were below the detection limit, the lowest measurable value of the assay was used for analysis.

### Statistical analyses

In a landmark study approach, Cox regression analyses were performed using time in years from date of follow-up assessment until occurrence of the primary end-point or census. To select a combination of clinical variables that best predicted outcomes, a least absolute shrinkage and selection operator (LASSO)-regularised Cox regression model was applied as implemented in the R-package glmnet (version 4.1.4). The regularisation parameter lambda was chosen from a grid of values through 10-fold cross-validation. To minimise overfitting, lambda was chosen by adding one standard error. Calibration of the model was assessed visually using calibration plots and by the Hosmer–Lemeshow goodness-of-fit test with  $p > 0.05$  indicating no significant deviation from perfect fit. Comparisons of area under the curves (AUCs) from receiver operating characteristic (ROC) were performed by the DeLong’s test utilising the R-package pROC (version 1.18.0). To improve comparability with the four-strata COMPERA 2.0 risk score that is recommended by the 2022 ESC/ERS guidelines on PH [1], we divided the newly derived model into quartiles based on all available data (*i.e.*, from baseline and follow-up assessments). We computed categorical net reclassification indices (NRI) indicating classification of patients into risk strata from different models as implemented in the R-package PredictABEL (version 1.2.4). The maximal Youden index for 1-year survival status from follow-up visits was used to compute optimal cut-offs. Kaplan–Meier survival curves with log-rank tests were performed by the R-package survival (version 3.4.0). Group comparisons were performed by chi-squared tests or Wilcoxon signed-rank tests as appropriate. Patients were randomly assigned to training or replication sub-cohorts using an in-built R-function (“sample”).

Missing data were imputed using the predictive mean modelling method as implemented in the R-package MICE (version 3.14.0). To increase the reliability, 25 imputed datasets were created (30 iterations each). Data were imputed in the range between minimal and maximal measured values. We included all clinical data but not the outcome variable in the imputation model. Results were averaged over all imputed datasets.

Data are presented as absolute numbers and percentages, or median and interquartile range (IQR). Results from regression analyses or ROC are presented along with 95% confidence intervals (CI). Correction for multiple testing was applied using the conservative Bonferroni method (q-values).

All statistical analyses were performed with R (version 4.2.1, R Foundation for Statistical Computing, Vienna, Austria).

## Results

### Cohort description

We included 298 patients with PAH (figure 1). Median age at diagnosis was 64 years (IQR 48–74), 61% were female and more than half of the patients were classified as idiopathic PAH (52%). Upfront combination therapy with targeted drugs was initiated in half of the patients (dual in 42% and triple in 8%), while the other half received initial monotherapy. Median time between treatment initiation and first comprehensive follow-up visit was 4.9 months (IQR 3.4–7.5), and during a median observational period of 3.5 years (IQR 1.7–5.9) from follow-up visit 119 composite outcome events occurred (114 deaths and five lung transplants). The rate of transplant-free survivors was 90% after 1 year, 72% after 3 years and 63% after 5 years from diagnosis and 86%, 71% and 62% from follow-up, respectively. The basic characteristics of patients are provided in table 1.

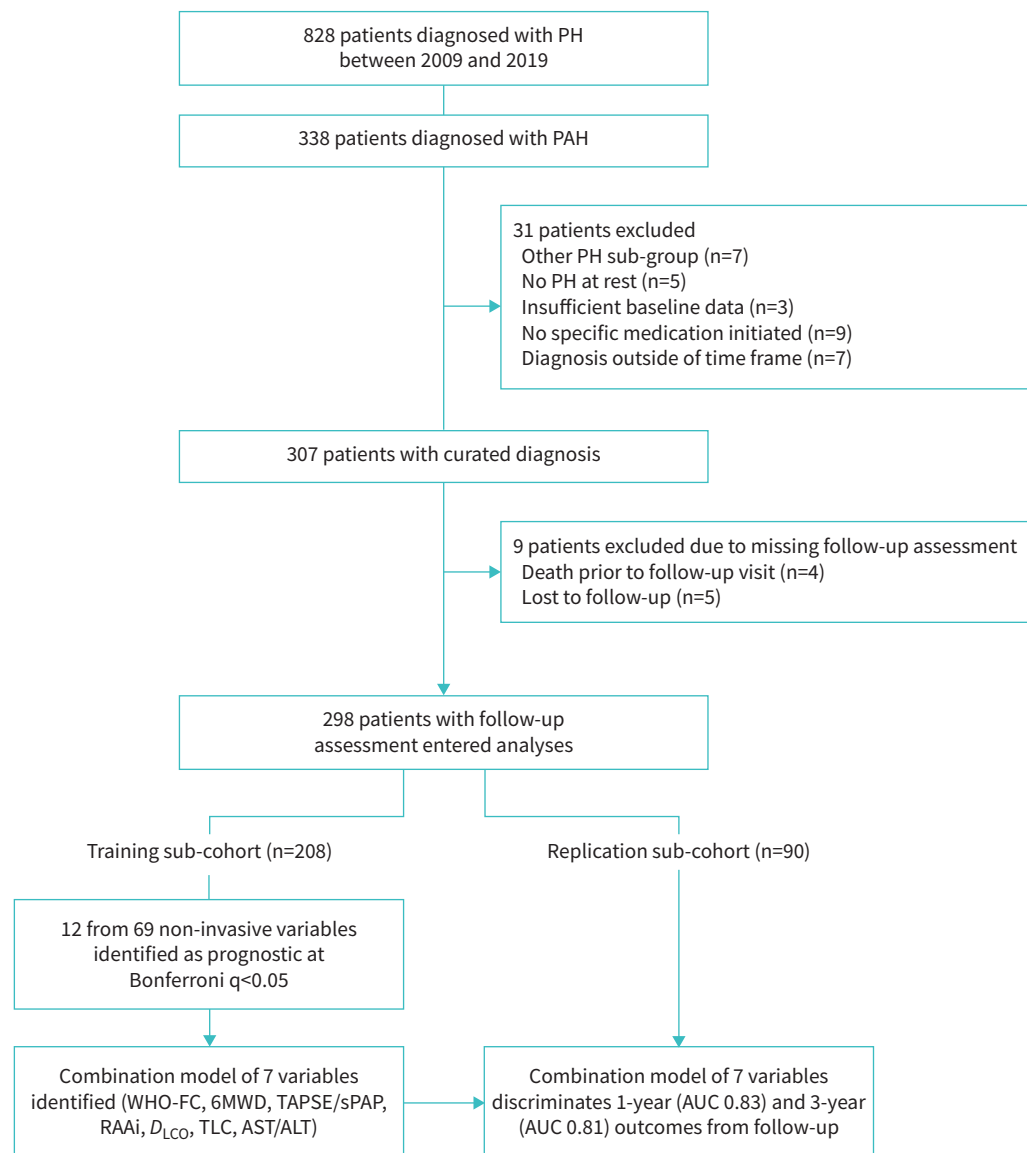
For subsequent analysis, we randomly separated our cohort in a 2:1 ratio into training (n=208) and replication (n=90) sub-cohorts. Both sub-cohorts showed comparable basic characteristics (table 1).

### Selection of outcome predictors from noninvasive follow-up assessment data

Using the training sub-cohort, we identified that 12 of 69 noninvasive clinical measures obtained at follow-up assessments were associated with outcomes in age- and sex-adjusted regular Cox regression models (Bonferroni-corrected  $q < 0.05$ ; table 2). All 12 entered LASSO-regularised modelling in the training sub-cohort, and seven variables were selected to create a single combinatory model (supplementary figure S1); namely, 6MWD, WHO-FC, tricuspid annular plane systolic excursion to systolic pulmonary arterial pressure ratio (TAPSE/sPAP), right atrial area index by the body surface area (RAAi), diffusing capacity of the lung for carbon monoxide ( $D_{LCO}$ ), total lung capacity (TLC) and aspartate transaminase to alanine transaminase ratio (AST/ALT; table 3).

### Internal replication and comparison with established risk models

The model based on seven noninvasive variables was calculated from follow-up assessment data in the internal replication sub-cohort and accurately discriminated 1-year and 3-year transplant-free survivors



**FIGURE 1** Study flowchart showing the selection of patients. 6MWD: 6-min walking distance; ALT: alanine aminotransferase; AST: aspartate aminotransferase; AUC: area under the curve;  $D_{\text{LCO}}$ : diffusing capacity of the lung for carbon monoxide; PH: pulmonary hypertension; PAH: pulmonary arterial hypertension; RAAi: right atrial area index; sPAP: systolic pulmonary arterial pressure; TAPSE: tricuspid annular plane systolic excursion; TLC: total lung capacity; WHO-FC: World Health Organization functional class.

from non-survivors after the follow-up visit with AUCs of 0.83 (95% CI 0.68–0.99,  $p=8.4\times 10^{-6}$ ) and 0.81 (95% CI 0.70–0.92,  $p=2.9\times 10^{-8}$ ), respectively (figure 2). The predicted and observed transplant-free survival rates matched closely indicating acceptable model calibration (supplementary figure S2), and the Hosmer–Lemeshow test indicated a good model fit for 1-year ( $p=0.891$ ) and 3-year ( $p=0.327$ ) outcomes.

The four-strata COMPERA 2.0 risk score also discriminated transplant-free survivors from non-survivors 1 and 3 years after follow-up in the replication sub-cohort (figure 2). Comparing the AUCs favoured the newly derived model for 1-year outcome status, both as continuous and as the four-strata version, but no significant difference occurred for 3-year outcome status (figure 2). In patients with an outcome event occurring within 1 year from follow-up, 33% (4 out of 12) were reclassified to a higher risk group based on the newly derived model; while 35% (25 out of 72) of patients without an outcome event within 1 year from follow-up entered a lower risk group (NRI for 1-year outcome status in the replication sub-cohort: 0.53, 95% CI 0.22–0.83,  $p<0.001$ ; supplementary table S2).

TABLE 1 Basic characteristics and treatment strategy of the study cohort

	All patients	Training sub-cohort	Replication sub-cohort
<b>Patients n</b>	298	208	90
<b>Demographics and disease classification</b>			
Age at diagnosis years	64 (48–74)	63 (46–74)	67 (53–74)
<b>Sex</b>			
Female	183 (61)	125 (60)	58 (64)
Male	115 (39)	83 (40)	32 (36)
BMI kg·m <sup>-2</sup>	25.8 (22.8–30.0)	26.0 (22.9–30.1)	25.7 (21.9–29.4)
<b>Disease classification</b>			
1.1	154 (52)	108 (52)	46 (51)
1.2	12 (4)	8 (4)	4 (4)
1.3	8 (3)	8 (4)	0 (0)
1.4.1	66 (22)	42 (20)	24 (27)
1.4.2	7 (2)	3 (1)	4 (4)
1.4.3	21 (7)	17 (8)	4 (4)
1.4.4	30 (10)	22 (11)	8 (9)
<b>Diagnostic haemodynamics</b>			
mPAP mmHg	43 (33–50)	43 (33–51)	41 (35–49)
mPAWP mmHg	12 (8–15)	11 (8–15)	12 (8–17)
PVR WU	6.4 (3.7–9.3)	6.8 (3.6–9.4)	5.6 (3.8–8.8)
Cardiac index L·min <sup>-1</sup> ·m <sup>-2</sup>	2.1 (1.7–2.6)	2.1 (1.7–2.6)	2.3 (1.7–2.6)
<b>Comorbidities</b>			
Hypertension	150 (50)	102 (49)	48 (53)
Coronary artery disease	81 (27)	55 (26)	26 (29)
Atrial fibrillation	67 (23)	40 (19)	27 (30)
Diabetes mellitus type II	59 (20)	40 (19)	19 (21)
COPD	51 (17)	32 (15)	19 (21)
<b>Initial treatment strategy</b>			
Monotherapy	150 (50)	104 (50)	46 (51)
Dual therapy	124 (42)	87 (42)	37 (41)
Triple therapy	24 (8)	17 (8)	7 (8)
<b>Outcome</b>			
<b>Events</b>			
All	119 (40)	85 (41)	34 (38)
Death	114 (38)	80 (38)	34 (38)
Lung transplant	5 (2)	5 (2)	0 (0)
Follow-up time years	3.5 (1.7–5.9)	3.1 (1.6–6.3)	4.0 (1.9–5.4)

Variables are presented as n (%) or as median (IQR) unless stated otherwise. Summary statistics are based on non-imputed data. BMI: body mass index; mPAP: mean pulmonary arterial pressure; mPAWP: mean pulmonary arterial wedge pressure; PVR: pulmonary vascular resistance; COPD: chronic obstructive pulmonary disease.

The noninvasive Registry to Evaluate Early And Long-term PAH Disease Management (REVEAL) 2.0 lite risk score also discriminated transplant-free survivors from non-survivors 1 or 3 years after follow-up in the replication sub-cohort (AUC 0.78 and 0.77). Comparisons of AUCs numerically favoured the newly derived model, albeit this was not statistically significant (DeLong's test  $p=0.151$  for AUC comparisons of 1-year outcome status and  $p=0.296$  for 3-year outcome status; supplementary figure S3).

#### Model sensitivity to aetiological subgroups of PAH

To assess the sensitivity of the newly derived model to aetiologies of PAH, we excluded patients with associated forms of PAH (APAH) using the entire cohort. The AUCs to discriminate transplant-free survivors from non-survivors 1 year from follow-up were 0.86 (95% CI 0.79–0.92,  $p=4.03\times 10^{-25}$ ,  $n=268$ ) following exclusion of patients with CHD-APAH and 0.90 (95% CI 0.84–0.96,  $p=2\times 10^{-40}$ ,  $n=174$ ) following exclusion of all patients with all forms of APAH. Similar results were obtained for the transplant-free survival status 3 years from follow-up with AUCs of 0.83 (95% CI 0.77–0.88,  $p=3.5\times 10^{-31}$ ) and 0.83 (95% CI 0.76–0.90,  $p=1.2\times 10^{-19}$ ), respectively. The four-strata version of the model based on quartiles likewise stratified outcomes from follow-up visits in Kaplan–Meier curves following the exclusion of patients with CHD-APAH and all APAH (figure 3).

TABLE 2 Comparison of prognostic variables identified

	All patients		Training sub-cohort		Replication sub-cohort	
	Baseline	Follow-up	Baseline	Follow-up	Baseline	Follow-up
<b>Patients n</b>	307	298	214	208	93	90
<b>Exercise capacity</b>						
WHO-FC						
I/II	52 (17)	111 (37)	32 (15)	72 (35)	20 (22)	39 (43)
III/IV	254 (83)	184 (62)	181 (85)	134 (64)	73 (79)	50 (56)
6MWD m	326 (237–420)	350 (270–460)	314 (233–411)	342 (270–426)	342 (263–424)	374 (276–483)
<b>Echocardiography</b>						
RAAi cm <sup>2</sup> ·m <sup>-2</sup> #	12.5 (10.3–15.2)	12.0 (9.7–15.1)	12.3 (10.0–15.4)	11.9 (10.0–15.0)	12.6 (10.6–15.2)	12.3 (9.2–15.4)
TAPSE/sPAP mm·mmHg <sup>-1</sup> #	0.242 (0.189–0.320)	0.285 (0.200–0.396)	0.231 (0.179–0.304)	0.290 (0.200–0.397)	0.259 (0.201–0.367)	0.278 (0.200–0.386)
<b>Pulmonary function</b>						
FEV <sub>1</sub> % pred	68 (56–82)	69 (55–81)	68 (58–81)	69 (55–81)	67 (55–83)	70 (55–81)
FVC % pred	72 (57–84)	71 (58–85)	71 (58–83)	71 (58–84)	72 (54–86)	71 (56–85)
TLC % pred	87 (77–98)	88 (77–98)	86 (76–97)	87 (76–98)	87 (79–98)	90 (77–99)
KCO % pred	60 (41–78)	60 (43–77)	60 (40–75)	60 (42–76)	62 (43–81)	61 (46–79)
D <sub>LCO</sub> % pred	50 (33–70)	53 (35–71)	51 (33–70)	54 (35–71)	50 (32–71)	52 (34–73)
<b>Biochemistry</b>						
NT-proBNP ng·L <sup>-1</sup>	1682 (441–4121)	876 (290–2368)	1782 (455–4019)	878 (309–2283)	1635 (355–4496)	793 (250–2449)
AST/ALT#	1.13 (0.89–1.53)	1.13 (0.88–1.48)	1.10 (0.88–1.54)	1.12 (0.87–1.49)	1.18 (0.91–1.50)	1.18 (0.92–1.48)
CRP ≥5 mg·L <sup>-1</sup>	141 (49)	129 (45)	102 (50)	95 (47)	39 (44)	34 (41)

Variables are presented as n (%) or as median (IQR) unless stated otherwise. There were no significant ( $p < 0.05$ ) differences between the training and replication sub-cohort for the presented variables. WHO-FC: World Health Organisation functional class; 6MWD: 6-min walking distance; RAAi: right atrial area indexed by the body surface area; TAPSE: tricuspid annular plane systolic excursion; sPAP: estimated systolic pulmonary arterial pressure by echocardiography; FEV<sub>1</sub>: forced expiratory volume in 1 s; % pred: percentage of the predicted reference value; FVC: forced vital capacity; TLC: total lung capacity; K<sub>CO</sub>: transfer coefficient of carbon monoxide; D<sub>LCO</sub>: diffusion capacity of the lung for carbon monoxide; NT-proBNP: N-terminal pro-brain natriuretic peptide; AST: aspartate aminotransferase; ALT: alanine aminotransferase; CRP: C-reactive protein. #: composite variables – summary statistics are based on non-imputed data except for composite variables, which were calculated after imputation.

### Model and individual component sensitivity to changes from baseline to follow-up assessment

We assessed changes of the newly derived model divided into four strata based on overall quartiles between baseline and follow-up assessments. After treatment initiation, 49% of patients entered a different stratum of the model at follow-up (35% down and 14% up; figure 4). In Kaplan–Meier survival analyses, these changes of risk strata translated into a tendency for outcome probability with better transplant-free survival for patients improving their risk stratum compared to those who entered a lower risk stratum (log-rank test on “improved” versus “worsened”  $p = 0.081$ ; supplementary figure S4).

Assessing each model component individually showed that out of the seven components, 6MWD, WHO-FC and TAPSE/sPAP significantly changed from baseline to follow-up (all  $p < 0.05$  for paired group comparisons; supplementary figure S5). 6MWD and WHO-FC, but not TAPSE/sPAP, were components of the COMPERA 2.0 score that is recommended by current European guidelines during follow-up in PAH [1, 12, 13]. Using an optimal cut-off value from Youden’s statistics ( $0.23 \text{ mm} \cdot \text{mmHg}^{-1}$ ; supplementary table 3), changes from baseline to follow-up in TAPSE/sPAP significantly translated into long-term outcomes in PAH ( $p = 0.003$  for log-rank test on “improved” versus “worsened”; figure 5).

### Adding TAPSE/sPAP to the four-strata COMPERA 2.0 risk score

All three components of the four-strata COMPERA 2.0 risk score including NT-proBNP were prognostic in our cohort (supplementary table S4). We next added TAPSE/sPAP to the COMPERA 2.0 risk score using cut-offs from the 2022 ESC/ERS guideline and that identified in our data:  $>0.32$  (low risk),  $0.24–0.32$  (intermediate–low risk),  $0.19–0.23$  (intermediate–high risk) and  $<0.19$  (high risk; supplementary table S5). In the total cohort, the addition of TAPSE/sPAP numerically increased the model’s discriminatory ability for 1-year (AUC from 0.80 to 0.81) and 3-year outcome status (AUC from 0.78 to 0.81), but this was not statistically significant (DeLong’s test  $p = 0.692$  and  $p = 0.110$ ).

### Discussion

Precise risk assessment is key to clinical management of patients with PAH, and sequential assessments based on noninvasive procedures promise to inform treatment decisions during the disease. We have used a

**TABLE 3** Results of the regular age- and sex-adjusted Cox regression analysis and  $\beta$ -coefficients of variables selected by LASSO modelling in the training sub-cohort

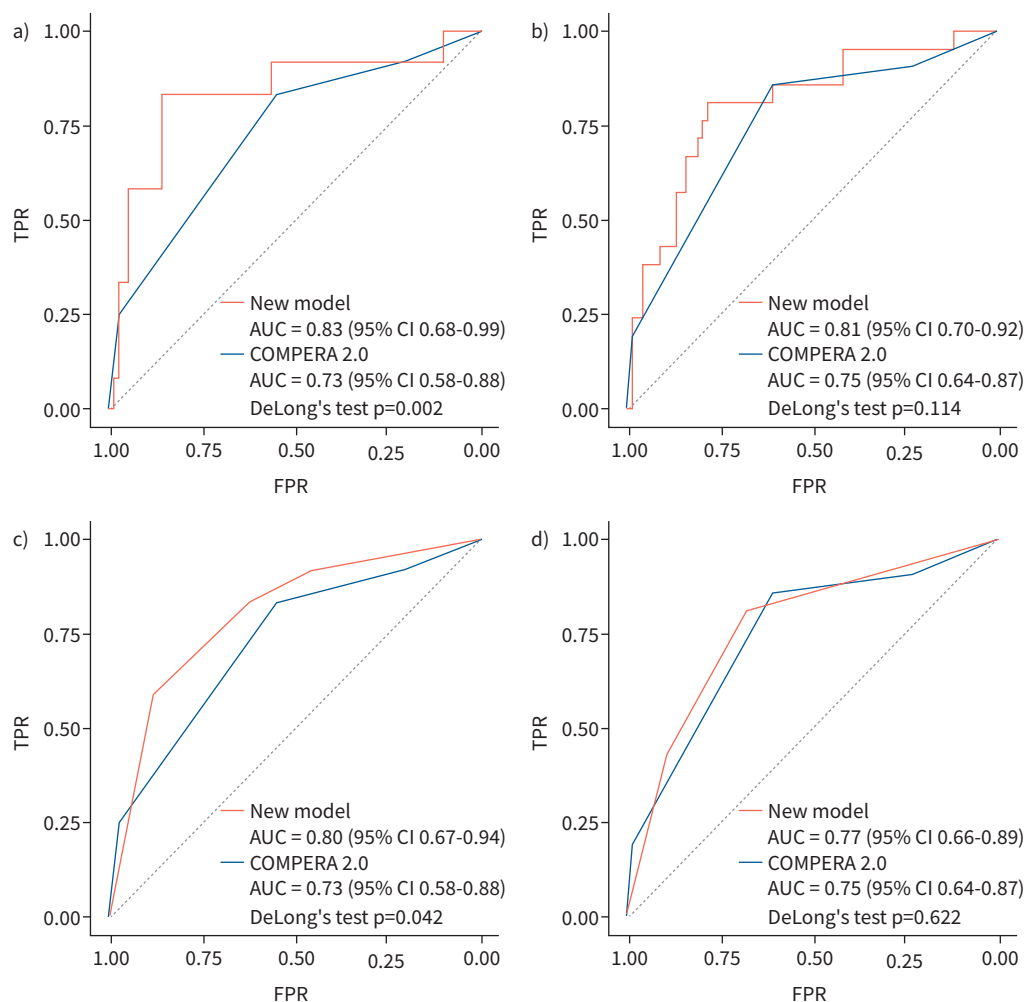
Variable	HR (95% CI)	p-value	q-value	LASSO $\beta$ -coefficient
<b>Prognostic variables selected by LASSO modelling</b>				
WHO-FC	3.123 (2.107–4.630)	<0.001	<0.001	0.420
6MWD m	0.995 (0.993–0.997)	<0.001	<0.001	–0.002
$D_{LCO}$ % pred	0.972 (0.960–0.984)	<0.001	<0.001	–0.009
AST/ALT	1.684 (1.322–2.145)	<0.001	0.002	0.110
RAAi $\text{cm}^2$	1.071 (1.031–1.112)	<0.001	0.024	0.006
TAPSE/sPAP $\text{mm}\cdot\text{mmHg}^{-1}$	0.038 (0.006–0.230)	<0.001	0.026	–0.069
TLC % pred	0.974 (0.960–0.989)	<0.001	0.033	–0.001
<b>Further prognostic variables (not selected by LASSO modelling)</b>				
$K_{CO}$ % pred	0.976 (0.965–0.986)	<0.001	<0.001	-
CRP $\text{mg}\cdot\text{L}^{-1}$	1.042 (1.023–1.062)	<0.001	0.001	-
FEV <sub>1</sub> % pred	0.977 (0.965–0.989)	<0.001	0.010	-
FVC % pred	0.979 (0.968–0.990)	<0.001	0.017	-
NT-proBNP $\text{ng}\cdot\text{L}^{-1}$	1.000 (1.000–1.000)	<0.001	0.049	-

Presented are hazard ratios (HR) with the respective 95% confidence intervals (95% CI) for the age- and sex-adjusted Cox regression analysis in the training sub-cohort from follow-up assessment data and the  $\beta$ -coefficients of the newly derived LASSO model for variables that were selected in the final model. The model intercept was zero. Summary statistics are based on imputed data. Bonferroni-adjustment was applied to obtain q-values. LASSO: least absolute shrinkage and selection operator; WHO-FC: World Health Organization functional class; 6MWD: 6-min walking distance;  $D_{LCO}$ : diffusion capacity of the lung for carbon monoxide; % pred: percentage of the predicted reference value; AST: aspartate aminotransferase; ALT: alanine aminotransferase; RAAi: right atrial area indexed by the body surface area; TAPSE: tricuspid annular plane systolic excursion; sPAP: estimated systolic pulmonary arterial pressure by echocardiography; TLC: total lung capacity;  $K_{CO}$ : transfer coefficient of carbon monoxide; CRP: C-reactive protein; FEV<sub>1</sub>: forced expiratory volume in 1 s; FVC: forced vital capacity; NT-proBNP: N-terminal pro-brain natriuretic peptide.

feature selection algorithm to combine noninvasive variables that best predicted future outcome events from follow-up visits in patients with PAH. A model combining seven variables well-stratified short-term mortality risk. Our systematic analyses indicate that residual risk can potentially be captured from noninvasive procedures during routine early follow-up visits in PAH and emphasises the need for extended data acquisition in large-scale registry studies.

Characteristics of our cohort including comorbidity seemed in general comparable to the two recent key studies on prognostic risk models from the two largest European registries on PAH, COMPERA and the French registry [12, 13]. Half of the patients received initial combination therapy, which was in line with recent data from the COMPERA registry, where the use of combination therapy increased from 2010 to 2019 to 46% during the first year of treatment [22]. COMPERA and our cohort contained predominantly cases with idiopathic disease [12]. In sensitivity analysis, we could not detect a reduction of the model's discriminatory ability following exclusion of associated forms of PAH. While numbers of patients in registries exceed vastly that from a single-centre study, single-centre data, through phenotypic depth, could pioneer variable selection.

The newly derived model achieved an AUC of 0.83 to discriminate 1-year outcome status in our unseen internal replication sub-cohort. The discriminative ability was comparable to Pulmonary Hypertension Outcomes Risk Assessment (PHORA) 2.0, which is based on a Bayesian network analysis in combination with random forest classification. In aggregated data from clinical trials on PAH, the noninvasive version of PHORA combining 12 variables achieved an AUC of 0.80 [23]. Interestingly, the addition of invasive measures only showed minimal increase, which underpins the potential of accurate risk prediction based on noninvasive clinical procedures. Similarly, in data from the French registry invasive haemodynamic variables were no longer significant following the introduction of BNP/NT-proBNP in the multivariate regression model [7]. Further biological markers that capture multiple pathophysiological processes, beyond myocardial stress, could enhance the biological component in risk models. Adding a six-protein score to NT-proBNP improved prediction of 5-year outcomes from AUC 0.76 to 0.82 in the combined UK National PAH Cohort Study and French Prospective Longitudinal Study of Patients With Idiopathic Pulmonary Arterial Hypertension, Family or Taking Anorectics (EFORT) [24]. Furthermore, imaging on the right heart can potentially refine risk assessment. The addition of a structural feature of the RV as

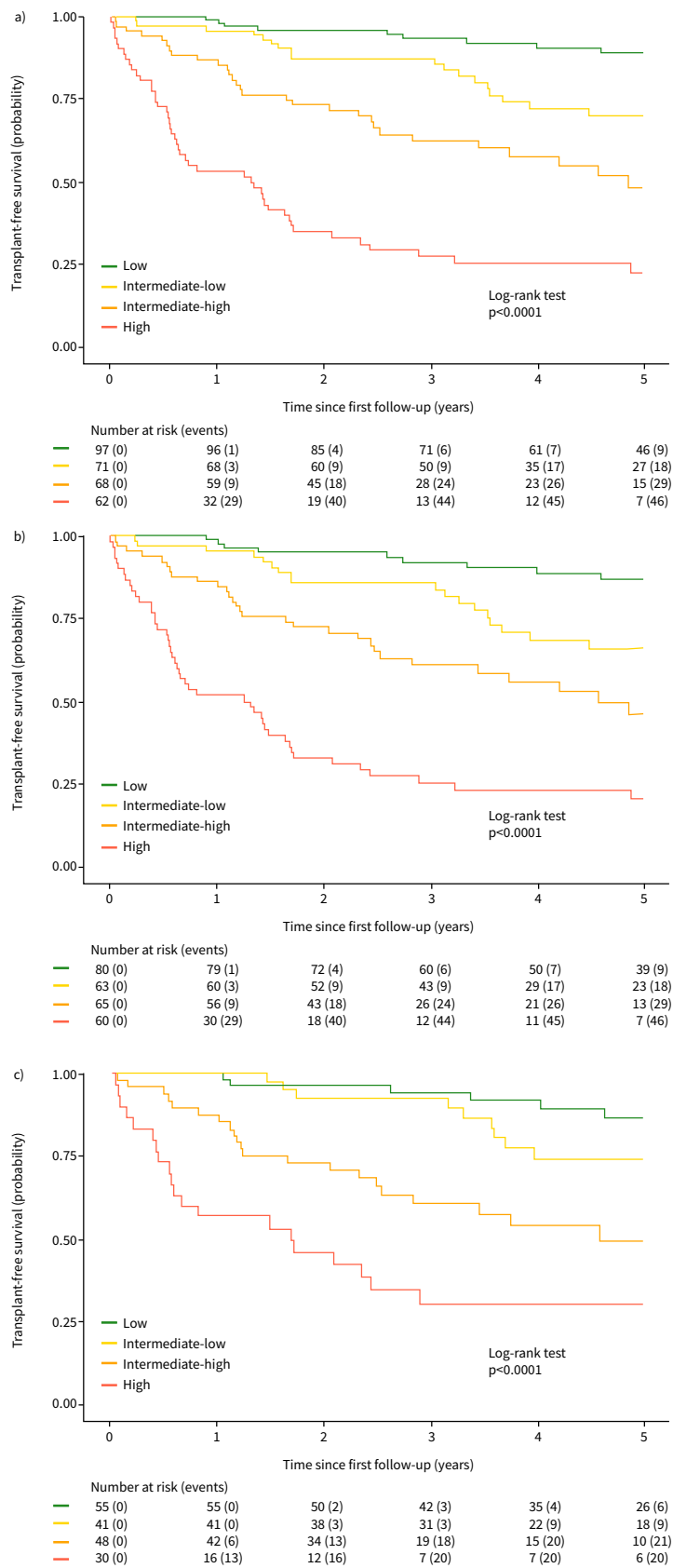


**FIGURE 2** Comparison of the areas under the curve (AUC) from receiver operating characteristic (ROC) with the corresponding 95% confidence intervals between the newly derived and COMPERA 2.0 model for the **a)** 1-year and **b)** 3-year outcomes after first follow-up assessment in the replication sub-cohort. ROC analyses were repeated using a four-strata approach to the newly derived model (based on quartiles) on **c)** 1-year and **d)** 3-year transplant-free survival. The grey dashed line marks the reference. TPR: true positive rate; FPR: false positive rate.

assessed by cardiac magnetic resonance (CMR) imaging (indexed RV end-systolic volume) to the REVEAL 2.0 risk score increased AUC from 0.74 to 0.78 for 1-year outcome status in a single-centre database [20]. However, advanced biological phenotyping and cardiac CMR are not readily available in all PH centres. Our approach, however, is entirely based on routine clinical variables that are time- and cost-effective.

The newly derived model combined seven noninvasive clinical variables with two overlapping with the COMPERA 2.0 score, WHO-FC and 6MWD [1], while NT-proBNP was not selected by LASSO modelling. For the reason of comparability and real-world applicability we used non-transformed values for each variable but cannot exclude that transforming the typically skewed distribution of NT-proBNP would have affected its selection. Additionally, (multi-)collinearity with other variables may have affected the selection of NT-proBNP. Next to exercise limitation as measured by WHO-FC and 6MWD, the seven variables in the LASSO model depict different pathophysiological aspects including impaired gas exchange ( $D_{LCO}$ ), restrictive ventilation pattern (TLC), impaired hepatic function (AST/ALT) and echocardiographic changes in right heart structure and function (RAAi, TAPSE/sPAP). Not all variables were sensitive to treatment initiation, when measurements were compared between baseline and follow-up, indicating possible association with non-modifiable PAH subgroups that intrinsically impact prognosis



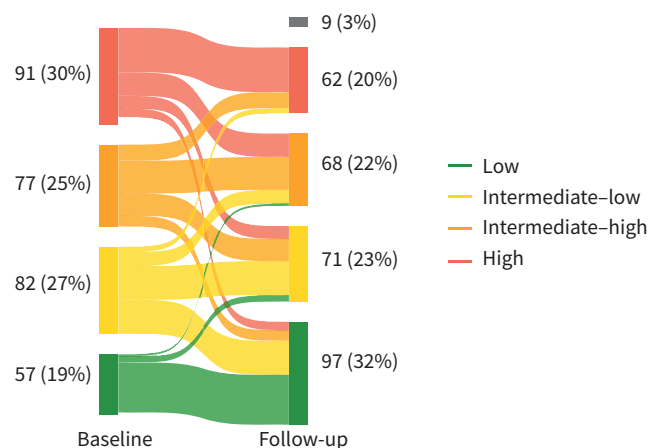


**FIGURE 3** Transplant-free survival from first follow-up assessment according to the newly derived model (divided into four-strata based on quartiles). **a)** For the entire cohort of patients with pulmonary arterial hypertension (PAH), **b)** after exclusion of patients with congenital heart disease-associated PAH and **c)** after exclusion of all associated forms of PAH.

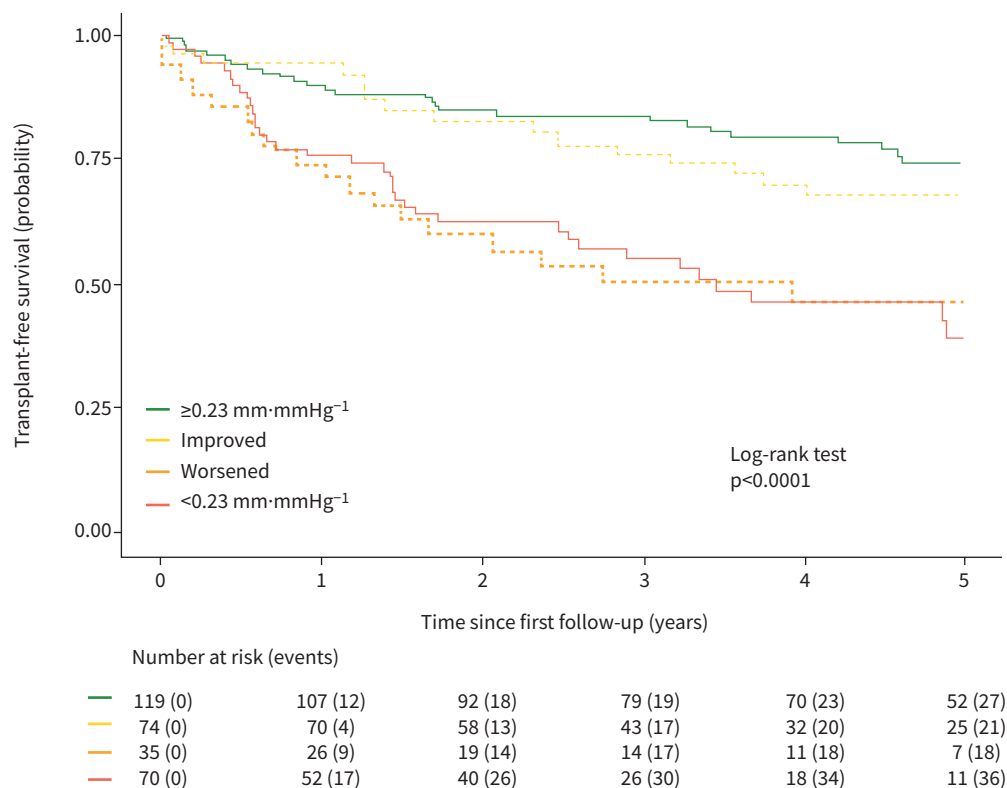
such as porto-pulmonary PH, CTD-PAH or the recently proposed lung phenotype in idiopathic PAH [25, 26]. But pulmonary vascular disease can also directly affect these variables. The liver can be affected by hepatic venous congestion and low cardiac output causing ischaemic hepatitis [5, 27].  $D_{LCO}$  can be moderately reduced possibly due to alveolar capillary membrane thickness and a reduction of pulmonary capillary blood volume caused by vascular remodelling [28, 29]. A mild restrictive pattern has been previously described as a common feature in PAH patients and is pronounced in patients with low  $D_{LCO}$  [26, 30]. Nevertheless, TLC was generally within the normal range in our cohort.

Although the 2015 and 2022 ESC/ERS guidelines recommend imaging of the right heart by echocardiography for risk assessment during diagnostic workup, sufficient evidence is missing on whether imaging can improve risk prediction during follow-up [1, 11]. It was shown that the RAA strongly correlates with the body surface area [31]. The indexed RAA, but not “raw” RAA, emerged from the discovery phase of our analysis, but it needs to be further established whether an improvement over RAA can be achieved by adjustment of the body composition. The ratio TAPSE/sPAP enables noninvasive estimation of the RV to pulmonary artery (PA) coupling [32, 33]. The TAPSE/sPAP showed an AUC of 0.83 to distinguish between RV-PA coupling and uncoupling as assessed by the gold-standard, pressure–volume relationships, with an optimal cut-off of  $0.31 \text{ mm}\cdot\text{mmHg}^{-1}$  [33]. We found an optimal cut-off for outcome discrimination of  $0.23 \text{ mm}\cdot\text{mmHg}^{-1}$ , which is at the lower end of those previously suggested ( $0.19\text{--}0.33 \text{ mm}\cdot\text{mmHg}^{-1}$ ) [1, 18, 33–36]. If TAPSE/sPAP increased during follow-up, patients showed significantly better long-term outcomes indicating its potential use to capture modifiable risk in PAH. Adding TAPSE/sPAP to the four-strata COMPERA 2.0 risk score increased numerically the AUCs for discrimination of outcome status.

Our study has limitations that should be discussed. We performed a single-centre study, sample size, geographical and ethnical diversity were limited, and we cannot exclude referral bias of more severe cases and cases with a pulmonary phenotype to our *tertiary* centre in respiratory medicine. Owing to the retrospective study design based on real-world patient management, there were missing data which were imputed to avoid selection bias and to enable statistical analysis, but we cannot exclude variance that was introduced by imputation. Some biochemical measures showed higher rates of missing data points but none of these were selected for model development. Review of original echocardiographic images was not possible in all cases. We only included procedures/variables that were available at our centre and



**FIGURE 4** Sankey diagram illustrating the change of risk status according to the newly derived model (divided into four strata based on quartiles) from baseline to follow-up. The width of the bars is proportional to the number of patients. The grey section marks the patients without follow-up assessment, which were excluded from development of the model.



**FIGURE 5** Kaplan–Meier survival plot of TAPSE to sPAP ratio from follow-up assessment. Variables were grouped by their optimal cut-offs ( $0.2343 \text{ mm} \cdot \text{mmHg}^{-1}$ ) and whether the variable improved or worsened as compared to baseline (*i.e.*, crossed the cut-off). Overall log-rank test is displayed. Comparisons of the curves for “improved” *versus* “worsened” were significant ( $p=0.003$  for log-rank test). sPAP: systolic pulmonary arterial pressure; TAPSE: tricuspid annular plane systolic excursion.

documented regularly during the study period. Thus, we were not able to include some predictive variables that were previously reported such as rate of hospitalisations before diagnosis of PAH [6]. External and ideally prospective validation on prevalent patients with PAH including different subgroups would be needed for this observational study to test for generalisability.

Our systematic approach indicates that incorporation of multi-modal routinely measured noninvasive clinical variables may increase accuracy of risk models in predicting future outcome events during follow-up in patients with PAH. We demonstrate that leveraging the phenotypic depth available at a single centre in combination with an unbiased feature selection algorithm can prioritise variables and inform large-scale registries and meta-analyses. Particularly easy to use variables obtained by echocardiographic imaging such as the TAPSE/sPAP ratio emerged as being dynamic and sensitive to changes in prognosis over time.

Provenance: Submitted article, peer reviewed.

Acknowledgement: The authors thank Anja Paulsen (University Medical Centre Hamburg-Eppendorf, Hamburg, Germany) for her work managing the data for the local PH registry. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sector.

Author contribution: Study design, data collection, statistical analyses and manuscript drafting were performed by J. Ostermann and L. Harbaum. All authors have contributed to data curation, interpretation of analyses and finalising the manuscript. All authors have approved the current version of the manuscript prior to submission and declared that the submitted work is original and has not been published (or being under consideration of publication) elsewhere (neither in English nor in any other language).

Conflict of interest: No financial conflict of interest has arisen from any author that interferes with the study design, methods or results.

## References

- 1 Humbert M, Kovacs G, Hoeper MM, *et al.* 2022 ESC/ERS Guidelines for the diagnosis and treatment of pulmonary hypertension. *Eur Respir J* 2023; 61: 2200879.
- 2 Humbert M, Guignabert C, Bonnet S, *et al.* Pathology and pathobiology of pulmonary hypertension: state of the art and research perspectives. *Eur Respir J* 2019; 53: 1801887.
- 3 D'Alonzo GE, Barst RJ, Ayres SM, *et al.* Survival in patients with primary pulmonary hypertension. Results from a national prospective registry. *Ann Intern Med* 1991; 115: 343–349.
- 4 Vonk Noordegraaf A, Chin KM, Haddad F, *et al.* Pathophysiology of the right ventricle and of the pulmonary circulation in pulmonary hypertension: an update. *Eur Respir J* 2019; 53: 1801900.
- 5 Rosenkranz S, Howard LS, Gomberg-Maitland M, *et al.* Systemic consequences of pulmonary hypertension and right-sided heart failure. *Circulation* 2020; 141: 678–693.
- 6 Benza RL, Gomberg-Maitland M, Miller DP, *et al.* The REVEAL Registry risk score calculator in patients newly diagnosed with pulmonary arterial hypertension. *Chest* 2012; 141: 354–362.
- 7 Boucly A, Weatherald J, Savale L, *et al.* Risk assessment, prognosis and guideline implementation in pulmonary arterial hypertension. *Eur Respir J* 2017; 50: 1700889.
- 8 Hoeper MM, Kramer T, Pan Z, *et al.* Mortality in pulmonary arterial hypertension: prediction by the 2015 European pulmonary hypertension guidelines risk stratification model. *Eur Respir J* 2017; 50: 1700740.
- 9 Benza RL, Miller DP, Gomberg-Maitland M, *et al.* Predicting survival in pulmonary arterial hypertension: insights from the Registry to Evaluate Early and Long-Term Pulmonary Arterial Hypertension Disease Management (REVEAL). *Circulation* 2010; 122: 164–172.
- 10 Kylhammar D, Kjellström B, Hjalmarsson C, *et al.* A comprehensive risk stratification at early follow-up determines prognosis in pulmonary arterial hypertension. *Eur Heart J* 2018; 39: 4175–4181.
- 11 Galiè N, Humbert M, Vachiery JL, *et al.* 2015 ESC/ERS Guidelines for the diagnosis and treatment of pulmonary hypertension: The Joint 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: Association for European Paediatric and Congenital Cardiology (AEPC), International Society for Heart and Lung Transplantation (ISHLT). *Eur Respir J* 2015; 46: 903–975.
- 12 Hoeper MM, Pausch C, Olsson KM, *et al.* COMPERA 2.0: a refined four-stratum risk assessment model for pulmonary arterial hypertension. *Eur Respir J* 2022; 60: 2102311.
- 13 Boucly A, Weatherald J, Savale L, *et al.* External validation of a refined four-stratum risk assessment score from the French pulmonary hypertension registry. *Eur Respir J* 2022; 59: 2102419.
- 14 Boucly A, Weatherald J, Humbert M, *et al.* Risk assessment in pulmonary arterial hypertension. *Eur Respir J* 2018; 51: 1800279.
- 15 Benza RL, Gomberg-Maitland M, Elliott CG, *et al.* Predicting survival in patients with pulmonary arterial hypertension: the REVEAL risk score calculator 2.0 and comparison with ESC/ERS-based risk assessment strategies. *Chest* 2019; 156: 323–337.
- 16 Benza RL, Kanwar MK, Raina A, *et al.* Development and validation of an abridged version of the REVEAL 2.0 risk score calculator, REVEAL Lite 2, for use in patients with pulmonary arterial hypertension. *Chest* 2021; 159: 337–346.
- 17 Olsson KM, Richter MJ, Kamp JC, *et al.* Refined risk stratification in pulmonary arterial hypertension and timing of lung transplantation. *Eur Respir J* 2022; 60: 2103087.
- 18 Fauvel C, Raitiere O, Boucly A, *et al.* Interest of TAPSE/SPAP ratio for noninvasive pulmonary arterial hypertension risk assessment. *J Heart Lung Transplant* 2022; 41: 1761–1772.
- 19 Harbaum L, Fuge J, Kamp JC, *et al.* Blood carbon dioxide tension and risk in pulmonary arterial hypertension. *Int J Cardiol* 2020; 318: 131–137.
- 20 Lewis RA, Johns CS, Cogliano M, *et al.* Identification of cardiac magnetic resonance imaging thresholds for risk stratification in pulmonary arterial hypertension. *Am J Respir Crit Care Med* 2020; 201: 458–468.
- 21 Kanwar MK, Gomberg-Maitland M, Hoeper M, *et al.* Risk stratification in pulmonary arterial hypertension using Bayesian analysis. *Eur Respir J* 2020; 56: 2000008.
- 22 Hoeper MM, Pausch C, Grünig E, *et al.* Temporal trends in pulmonary arterial hypertension: results from the COMPERA registry. *Eur Respir J* 2022; 59: 2102024.
- 23 Fauvel C, Liu Z, Lin S, *et al.* Pulmonary arterial hypertension risk assessment model using Bayesian Network analysis. *Eur Respir J* 2022; 60: Suppl. 66, 2137.
- 24 Rhodes CJ, Wharton J, Swietlik EM, *et al.* Using the plasma proteome for risk stratifying patients with pulmonary arterial hypertension. *Am J Respir Crit Care Med* 2022; 205: 1102–1111.
- 25 Gall H, Felix JF, Schneck FK, *et al.* The Giessen Pulmonary Hypertension Registry: survival in pulmonary hypertension subgroups. *J Heart Lung Transplant* 2017; 36: 957–967.
- 26 Hoeper MM, Dwivedi K, Pausch C, *et al.* Phenotyping of idiopathic pulmonary arterial hypertension: a registry analysis. *Lancet Respir Med* 2022; 10: 937–948.
- 27 Yogeswaran A, Tello K, Lund J, *et al.* Risk assessment in pulmonary hypertension based on routinely measured laboratory parameters. *J Heart Lung Transplant* 2022; 41: 400–410.

- 28 Oppenheimer BW, Berger KI, Hadjiangelis NP, *et al.* Membrane diffusion in diseases of the pulmonary vasculature. *Respir Med* 2006; 100: 1247–1253.
- 29 Farha S, Laskowski D, George D, *et al.* Loss of alveolar membrane diffusing capacity and pulmonary capillary blood volume in pulmonary arterial hypertension. *Respir Res* 2013; 14: 6.
- 30 Sun XG, Hansen JE, Oudiz RJ, *et al.* Pulmonary function in primary pulmonary hypertension. *J Am Coll Cardiol* 2003; 41: 1028–1035.
- 31 Grünig E, Henn P, D'Andrea A, *et al.* Reference values for and determinants of right atrial area in healthy adults by 2-dimensional echocardiography. *Circ Cardiovasc Imaging* 2013; 6: 117–124.
- 32 Guazzi M, Bandera F, Pelissero G, *et al.* Tricuspid annular plane systolic excursion and pulmonary arterial systolic pressure relationship in heart failure: an index of right ventricular contractile function and prognosis. *Am J Physiol Heart Circ Physiol* 2013; 305: H1373–H1381.
- 33 Tello K, Wan J, Dalmer A, *et al.* Validation of the tricuspid annular plane systolic excursion/systolic pulmonary artery pressure ratio for the assessment of right ventricular-arterial coupling in severe pulmonary hypertension. *Circ Cardiovasc Imaging* 2019; 12: e009047.
- 34 Tello K, Axmann J, Ghofrani HA, *et al.* Relevance of the TAPSE/PASP ratio in pulmonary arterial hypertension. *Int J Cardiol* 2018; 266: 229–235.
- 35 Kazimierczyk R, Kazimierczyk E, Knapp M, *et al.* Echocardiographic assessment of right ventricular-arterial coupling in predicting prognosis of pulmonary arterial hypertension patients. *J Clin Med* 2021; 10: 2995.
- 36 Vicenzi M, Caravita S, Rota I, *et al.* The added value of right ventricular function normalized for afterload to improve risk stratification of patients with pulmonary arterial hypertension. *PLoS One* 2022; 17: e0265059.