



External validation of a refined four-stratum risk assessment score from the French pulmonary hypertension registry

Athénaïs Boucly ^{1,2,3,13}, Jason Weatherald ^{4,13}, Laurent Savale ^{1,2,3}, Pascal de Groot⁵, Vincent Cottin ⁶, Grégoire Prévot⁷, Ari Chaouat ⁸, François Picard⁹, Delphine Horeau-Langlard¹⁰, Arnaud Bourdin ¹¹, Etienne-Marie Jutant ^{1,2,3,12}, Antoine Beurnier ^{1,2,3}, Mitja Jevnikar^{1,2,3}, Xavier Jaïs^{1,2,3}, Gérald Simonneau^{1,2,3}, David Montani ^{1,2,3}, Olivier Sitbon ^{1,2,3,14} and Marc Humbert ^{1,2,3,14}

¹Université Paris-Saclay, School of Medicine, Le Kremlin-Bicêtre, France. ²Dept of Respiratory and Intensive Care Medicine, AP-HP, Hôpital Bicêtre, Le Kremlin-Bicêtre, France. ³INSERM UMR_S 999, Hôpital Marie Lannelongue, Le Plessis-Robinson, France. ⁴University of Calgary, Dept of Medicine, Division of Respiratory, and Libin Cardiovascular Institute, Calgary, AB, Canada. ⁵Université de Lille, Service de Cardiologie, CHU Lille, Institut Pasteur de Lille, Inserm U1167, Lille, France. ⁶Université Lyon 1, INRAE, UMR754, IVPC, National Reference Centre for Rare Pulmonary Diseases, Hospices Civils de Lyon, Lyon, France. ⁷CHU de Toulouse, Hôpital Larrey, Service de Pneumologie, Toulouse, France. ⁸Inserm UMR_S1116, Faculté de Médecine de Nancy, Université de Lorraine, Département de Pneumologie, CHRU de Nancy, Vandoeuvre-lès-Nancy, France. ⁹Université Bordeaux, Hôpital Cardiologique du Haut-Lévêque, Heart Failure Unit and Pulmonary Hypertension Expert Centre, Bordeaux, France. ¹⁰CHU de Nantes, Hôpital Laënnec, Service de Pneumologie, Nantes, France. ¹¹Université Montpellier, CHU Montpellier, Dept of Respiratory Diseases, Montpellier, France. ¹²CHU Poitiers, Service de Pneumologie, Poitiers, France. ¹³Both authors contributed equally. ¹⁴Both authors contributed equally.

Corresponding author: Marc Humbert (marc.humbert@aphp.fr)



Shareable abstract (@ERSpublications)

A four-stratum risk assessment method with low, intermediate-low, intermediate-high and high risk categories was better at discriminating survival in pulmonary arterial hypertension than a three-stratum method with low, intermediate and high risk groups <https://bit.ly/3mA6kj7>

Cite this article as: 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 [DOI: 10.1183/13993003.02419-2021].

Copyright ©The authors 2022.

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

Received: 6 Sept 2021
Accepted: 28 Oct 2021

Abstract

Introduction Contemporary risk assessment tools categorise patients with pulmonary arterial hypertension (PAH) as low, intermediate or high risk. A minority of patients achieve low risk status with most remaining intermediate risk. Our aim was to validate a four-stratum risk assessment approach categorising patients as low, intermediate-low, intermediate-high or high risk, as proposed by the Comparative, Prospective Registry of Newly Initiated Therapies for Pulmonary Hypertension (COMPERA) investigators.

Methods We evaluated incident patients from the French PAH Registry and applied a four-stratum risk method at baseline and at first reassessment. We applied refined cut-points for three variables: World Health Organization functional class, 6-min walk distance and N-terminal pro-brain natriuretic peptide. We used Kaplan–Meier survival analyses and Cox proportional hazards regression to assess survival according to three-stratum and four-stratum risk approaches.

Results At baseline (n=2879), the four-stratum approach identified four distinct risk groups and performed slightly better than a three-stratum method for predicting mortality. Four-stratum model discrimination was significantly higher than the three-stratum method when applied during follow-up and refined risk categories among subgroups with idiopathic PAH, connective tissue disease-associated PAH, congenital heart disease and portopulmonary hypertension. Using the four-stratum approach, 53% of patients changed risk category from baseline compared to 39% of patients when applying the three-stratum approach. Those who achieved or maintained a low risk status had the best survival, whereas there were more nuanced differences in survival for patients who were intermediate-low and intermediate-high risk.

Conclusions The four-stratum risk assessment method refined risk prediction, especially within the intermediate risk category of patients, performed better at predicting survival and was more sensitive to change than the three-stratum approach.



Introduction

In 2015, the European Society of Cardiology (ESC) and European Respiratory Society (ERS) guidelines proposed a multidimensional risk stratification tool to guide prognostication and treatment decisions for patients with pulmonary arterial hypertension (PAH) [1]. The 2015 ESC/ERS guidelines recommended categorisation of patients into low (<5% estimated risk of 1-year mortality), intermediate (5–10% estimated risk of 1-year mortality) and high risk (>10% estimated 1-year mortality) using clinical, exercise, imaging and haemodynamic variables known to be associated with prognosis [1].

Shortly after the 2015 ESC/ERS guidelines were published, several registry-based studies from Europe proposed methods of implementing this risk assessment proposal [2–5]. The Swedish Pulmonary Arterial Hypertension Register (SPAHR) and the Comparative, Prospective Registry of Newly Initiated Therapies for Pulmonary Hypertension (COMPERA) group used an integer score method which assigned values of 1, 2 or 3 to each variable corresponding to their low, intermediate or high risk cut-points in the 2015 ESC/ERS guidelines risk table. They then calculated the average value for each patient [2, 3]. Similar to SPAHR/COMPERA, the French Pulmonary Hypertension (PH) Registry approach included clinical, exercise and invasive haemodynamic variables, but the French approach differed in methodology. Instead of an integer score, the French PH registry method counted the number of variables meeting the low risk criteria definition at baseline and first follow-up for World Health Organization (WHO) functional class (FC), 6-min walk distance (6MWD), right atrial pressure and cardiac index [4, 5]. A simplified noninvasive French PH Registry approach using only three noninvasive low risk variables (6MWD >440 m, WHO FC I or II and N-terminal pro-brain natriuretic peptide (NT-proBNP) <300 ng·L⁻¹ or BNP <50 ng·L⁻¹) can identify a truly low risk group of patients with 1- and 5-year survival ≥95% [4, 6].

The SPAHR, COMPERA and French PH Registry scores use overlapping variables and cut-points with the United States Registry to Evaluate Early and Long-Term PAH Disease Management (REVEAL) score, which also includes nonmodifiable prognostic factors such as disease aetiology, age and sex [7]. The updated REVEAL 2.0 score classifies patients similarly into three categories (low, intermediate and high risk) with corresponding 1-year mortality estimates of 1.9%, 6.5% and 25.8% [8]. These approaches have also been validated in *post hoc* analyses of the PATENT-1 trial of riociguat in PAH [9, 10]. An abridged version of the REVEAL 2.0 score, REVEAL 2.0 Lite, uses six modifiable variables and revised cut-points for noninvasive variables (WHO FC, systolic blood pressure, heart rate, 6MWD and NT-proBNP/BNP) [11].

Still, advances in risk stratification are needed. Discrimination characteristics of the SPAHR/COMPERA, French PH Registry and REVEAL 2.0 scores are good, but not excellent, and could be further improved [8, 12, 13]. Furthermore, it remains uncertain what the best treatment strategy is for patients who remain in the intermediate risk group using a three-stratum approach. In the European registry studies, a minority of patients achieved a low risk profile with initial PAH treatment and the majority of patients were in the intermediate risk category at baseline and during follow-up [2–4]. Thus, a more nuanced approach with more refined definition of intermediate risk patients may help better inform treatment decisions. To address this problem of the intermediate risk group, the SPAHR investigators suggested subdividing the intermediate risk group into intermediate-low risk and intermediate-high risk [14]. A four-stratum risk approach described by the COMPERA investigators using revised scoring and cut-points for the 6MWD, WHO FC and NT-proBNP/BNP may better define risk groups [15]. The objective of this study was to validate this approach by assessing whether a four-stratum risk assessment strategy is associated with survival among patients with PAH from the French PH registry.

Methods

Study design

This was a retrospective analysis of prospectively collected data in the French PH registry (<http://registre-htap.aphp.fr>). Although French law does not require ethics committee approval or informed consent for retrospective data collection, the data were anonymised and compiled according to the requirements of the organisation dedicated to privacy, information technology and civil rights in France (“CNIL”). The committee approved the methods used to collect and analyse data on 24 May 2003 (approval number 842063). The current study complied with the Declaration of Helsinki. The French PH registry is part of the French Pulmonary Hypertension Reference Center (PulmoTension), funded by the French Ministry of Health.

Patient population

Data were collected using the web-based PAHTool platform (Inovultus, Santa Maria da Feira, Portugal). We reviewed data from all incident patients with group 1 PAH who were enrolled in the French PH Registry between 1 January 2009 and 31 December 2020. Inclusion criteria were 1) adults (aged

≥18 years); 2) a right heart catheterisation (RHC) demonstrating pre-capillary PAH, defined as mean pulmonary arterial pressure (mPAP) ≥25 mmHg, pulmonary arterial wedge pressure ≤15 mmHg and a pulmonary vascular resistance >3 Wood units. Patients were excluded if they had known pulmonary veno-occlusive disease, unrepaired congenital heart disease patients, including those with Eisenmenger syndrome, or were missing data for WHO FC, 6MWD and/or NT-proBNP/BNP at baseline.

Risk stratification

Patients were classified using the three-stratum SPAHR/COMPERA approach (low, intermediate, high) as described previously [2, 3], as well as with a four-stratum model using cut-points for WHO FC, 6MWD and NT-proBNP/BNP (table 1) based on cut-points derived and used in the COMPERA 2.0 analysis [15]. A score of 1 was assigned for low risk, 2 for intermediate-low risk, 3 for intermediate-high risk, and 4 for high risk values, then an average was calculated for each patient, rounded to the nearest integer. Thus, an average score of <1.5 classified a patient as low risk, a score of 1.5–2.49 was intermediate-low risk, 2.5–3.49 was intermediate-high risk and ≥3.5 was classified as high risk. We assessed overall survival according the four-stratum score at baseline and at the time of first follow-up within 3–24 months after diagnosis.

Statistical analysis

Continuous data are represented as mean±SD or median (interquartile range (IQR), 25–75%) according to data distribution. Categorical data are expressed as number (n) and percentage (%). The primary outcome was all-cause mortality. Survival time was calculated from the date of diagnostic RHC until death or last recorded clinical contact. Patients who underwent lung transplantation were censored on the date of transplantation. Survival analyses were performed using the Kaplan–Meier method with the log rank test. Cox proportional hazards regression was used to assess the association between risk category and survival, expressed as hazard ratios with 95% confidence intervals. We used Harrell's C-statistic and the Akaike information criterion (AIC) to compare model goodness of fit of the Cox model for discriminating overall and 1-year mortality for the three-stratum and four-stratum risk methods. We compared model performance using Harrell's C and Somers' D by randomly splitting the cohort into training and test sets [16]. Statistical significance was set at $\alpha \leq 0.05$. Statistical analysis was performed using SPSS Statistics (version 26; SPSS, Chicago, IL, USA) and STATA (version 13.1; StataCorp, College Station, TX, USA).

Results

Patient characteristics

Among 4382 newly diagnosed patients with PAH enrolled in the French PH registry between 1 January 2009 and 31 December 2020, 2879 patients met eligibility criteria and were included (figure 1). There were 2082 patients with available data for a follow-up risk reassessment. Characteristics at baseline according to the four risk strata are shown in table 2. The mean age was 61±15 years and 60% were female. Idiopathic PAH was the most frequent aetiology (38%), followed by connective tissue disease (CTD)-associated PAH (27%). The median observation time was 2.25 years (IQR 0.71–4.57 years) and 1092 (38%) patients died during the follow-up period. The overall 1-, 3- and 5-year survival in this cohort was 88%, 69% and 52%, respectively (supplementary figure E1).

Risk assessment at baseline

Using the three-stratum SPAHR/COMPERA risk assessment method, most patients (67%) were classified as intermediate risk at baseline, with 16% classified as low risk and 16% classified as high risk. Using the four-stratum approach, 12% were low risk, 40% were intermediate-low risk, 33% were intermediate-high

TABLE 1 Proposed scoring for the Comparative, Prospective Registry of Newly Initiated Therapies for Pulmonary Hypertension (COMPERA) 2.0 four-stratum risk assessment method

	Points assigned			
	1	2	3	4
WHO FC	I or II		III	IV
6MWD, m	>440	440–320	319–165	<165
BNP, ng·L ⁻¹	<50	50–199	200–800	>800
NT-proBNP, ng·L ⁻¹	<300	300–649	650–1100	>1100

WHO: World Health Organization; FC: functional class; 6MWD: 6-min walk distance; BNP: brain natriuretic peptide, NT-proBNP: N-terminal pro-BNP.

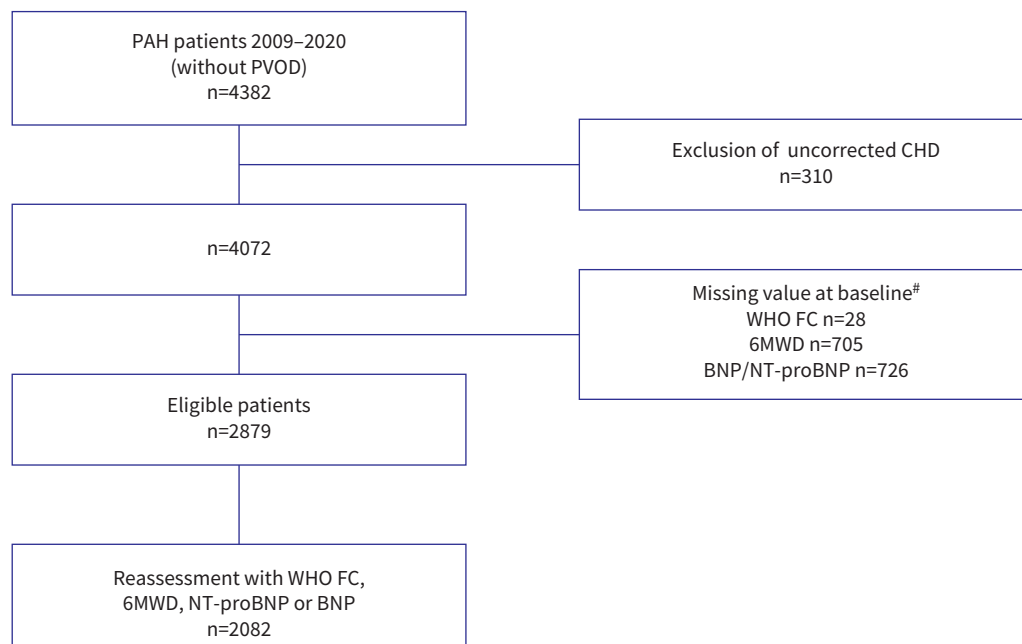


FIGURE 1 Study flow diagram. PAH: pulmonary arterial hypertension; PVOD: pulmonary veno-occlusive disease; CHD: congenital heart disease; WHO FC: World Health Organization functional class; 6MWD: 6-min walk distance; BNP: brain natriuretic peptide; NT-proBNP: N-terminal pro-BNP. #: more than one reason for exclusion could apply.

risk and 15% were high risk. Overall survival from diagnosis using the three-stratum and four-stratum risk scores is shown in figure 2. There were significant differences in survival across risk groups using the three-stratum approach and using the four-stratum approach. Using the four risk strata at baseline, the low risk group had an estimated 1-, 3- and 5-year survival of 98%, 89% and 75%, respectively. For the intermediate-low risk group, 1-, 3- and 5-year survival was 93%, 81% and 65%, respectively. For the intermediate-high risk group, 1-, 3- and 5-year survival was 86%, 63% and 44%, respectively. For the high risk group, 1-, 3- and 5-year survival was 75%, 45% and 31%, respectively.

In Cox proportional hazards regression models, there was an increasing risk of death for patients in the intermediate and high risk groups compared to low risk groups at baseline using both stratification methods (supplementary table E1). The four-stratum model discrimination for overall mortality was slightly higher (Harrell's C-statistic 0.64, AIC 15238.4) than the three-stratum method (Harrell's C 0.61, AIC 15296.0), but this was not significantly different ($p>0.05$). The four-stratum model discrimination for 1-year mortality after diagnosis was also modestly, but significantly higher compared to the three-stratum model (Harrell's C 0.67, AIC 4470.4 versus Harrell's C 0.63, AIC 4500.9; $p<0.001$).

Risk assessment at follow-up

There were 2082 patients with complete data to calculate a three-stratum and four-stratum risk score at follow-up. The median duration between diagnosis and first reassessment for this analysis was 5.1 months (IQR 3.9–9.7 months). Using the three-stratum method, 39% were low risk, 53% were intermediate risk and 8% were high risk at the time of first reassessment. Using the four-stratum method, 33% were classified as low risk, 38% were intermediate-low risk, 23% were intermediate-high risk and 6% were high risk. Overall survival after first reassessment according to the three-stratum and four-stratum models is shown in figure 2. Using the four risk strata at first reassessment, the low risk group had an estimated 1-, 3- and 5-year survival of 97%, 89% and 81%, respectively. For the intermediate-low risk group, 1-, 3- and 5-year survival was 94%, 75% and 57%, respectively. For the intermediate-high risk group 1-, 3- and 5-year survival was 81%, 50% and 31%, respectively. For the high risk group, 1-, 3- and 5-year survival was 65%, 28% and 13%, respectively.

In Cox regression models, there was increased risk of mortality after the first reassessment with increasing risk strata (supplementary table E1). Similar to the baseline risk assessment, the four-stratum

TABLE 2 Baseline characteristics

	All patients	Low risk	Intermediate-low risk	Intermediate-high risk	High risk
Patients, n	2879	340	951	1162	426
Age, years	61±15	54±14	59±14	63±14	65±15
Female	1737 (60)	181 (53)	562 (59)	720 (62)	274 (64)
BMI, kg·m⁻²	27.2±6.5	26.1±4.9	27.4±6.2	27.6±7.0	26.6±6.7
Aetiology of PAH					
Idiopathic	1094 (38)	99 (29)	323 (34)	483 (41.5)	189 (44)
Heritable	137 (5)	23 (7)	45 (5)	56 (5)	13 (3)
Drug- and toxin-induced	230 (8)	24 (7)	77 (8)	102 (9)	27 (6)
CTD	781 (27)	93 (27)	244 (26)	297 (26)	147 (35)
SSc	603 (21)	71 (21)	188 (20)	236 (20)	108 (25)
CHD	23 (1)	6 (2)	7 (0.5)	10 (1)	0
HIV	89 (3)	19 (6)	38 (4)	28 (2)	4 (1)
PoPH	525 (18)	76 (22)	217 (23)	186 (16)	46 (11)
Comorbidities					
Obesity	654 (23)	56 (16)	221 (23)	284 (24)	93 (22)
Coronary heart disease	182 (6)	15 (4)	43 (5)	87 (7)	37 (9)
Diabetes mellitus	487 (17)	32 (9)	148 (16)	212 (18)	95 (22)
Arterial hypertension	1222 (42)	95 (28)	377 (40)	549 (47)	201 (47)
WHO FC					
I–II	925 (32)	340 (100)	487 (51)	98 (8)	0
III	1541 (54)	0	456 (48)	925 (80)	160 (38)
IV	413 (14)	0	8 (1)	139 (12)	266 (62)
6MWD, m	300 (176–400)	466 (420–513)	367 (306–426)	248 (180–325)	0 (0–115)
NT-proBNP, ng·L⁻¹	995 (281–2726)	135 (78–247)	422 (161–858)	1573 (777–3020)	3597 (194–7074)
BNP, ng·L⁻¹	207 (74–512)	35 (20–61)	108 (50–225)	360 (177–616)	880 (499–1286)
Haemodynamics					
RAP, mmHg	8±5	6±4	7±5	9±6	11±7
mPAP, mmHg	45±12	39±12	43±12	47±12	49±12
PAWP, mmHg	9±4	9±4	9±4	9±4	9±4
Cardiac output, L·min ⁻¹	4.6±1.6	5.6±1.5	5.1±1.6	4.3±1.4	3.8±1.2
Cardiac index, L·min ⁻¹ ·m ⁻²	2.6±0.8	3.1±0.8	2.8±0.8	2.4±0.7	2.2±0.6
PVR, Wood units	8.8±4.8	5.8±2.9	7.4±4.1	9.7±4.8	11.6±5.8
S _{vO₂} , %	63±10	71±7	66±7	61±9	56±12
Heart rate, beats·min ⁻¹	79±15	75±14	76±15	79±16	84±16
SVI, mL·m ⁻²	35±17	45±30	39±13	31±10	26±7
Initial treatment strategy					
CCB only	167 (6)	37 (11)	59 (6)	62 (5)	9 (2)
Monotherapy	1397 (48)	193 (57)	530 (56)	531 (46)	143 (34)
Dual therapy	796 (28)	51 (15)	204 (21)	387 (33)	154 (36)
Triple therapy	79 (3)	1 (0.5)	14 (2)	35 (3)	29 (7)
None	440 (15)	58 (17)	144 (15)	147 (13)	91 (21)

Data are presented as mean±sd, n (%) or median (interquartile range), unless otherwise stated. BMI: body mass index; PAH: pulmonary arterial hypertension; CTD: connective tissue disease; SSc: systemic sclerosis; CHD: congenital heart disease; PoPH: portopulmonary hypertension; WHO FC: World Health Organization functional class; 6MWD: 6-min walk distance; NT-proBNP: N-terminal pro-brain natriuretic peptide; RAP: right atrial pressure; mPAP: mean pulmonary arterial pressure; PAWP: pulmonary artery wedge pressure; PVR: pulmonary vascular resistance; S_{vO₂}: mixed venous oxygen saturation; SVI: stroke volume index; CCB: calcium channel blocker.

discrimination for overall mortality after first reassessment was slightly, but significantly, higher compared to the three-stratum method (Harrell's C 0.70, AIC 9242.7 versus Harrell's C 0.67, AIC 9299.3; p<0.001). The four-stratum model discrimination was also higher for 1-year mortality after the first reassessment compared to the three-stratum model (Harrell's C 0.73, AIC 2434.8, versus Harrell's C 0.69, AIC 2466.0; p=0.001). Given that lung transplant may be a competing risk for death in eligible patients we performed a competing risk analysis, which did not change the results at baseline or follow-up (data not shown).

Changes in four-stratum risk assessment

In the overall population (n=2879) we used Sankey diagrams to represent changes in risk category using the three-stratum and four-stratum methods (figure 3). According to the three-stratum method, there was an increase in the proportion of patients in the low risk category from baseline (16%) to follow-up (28%).

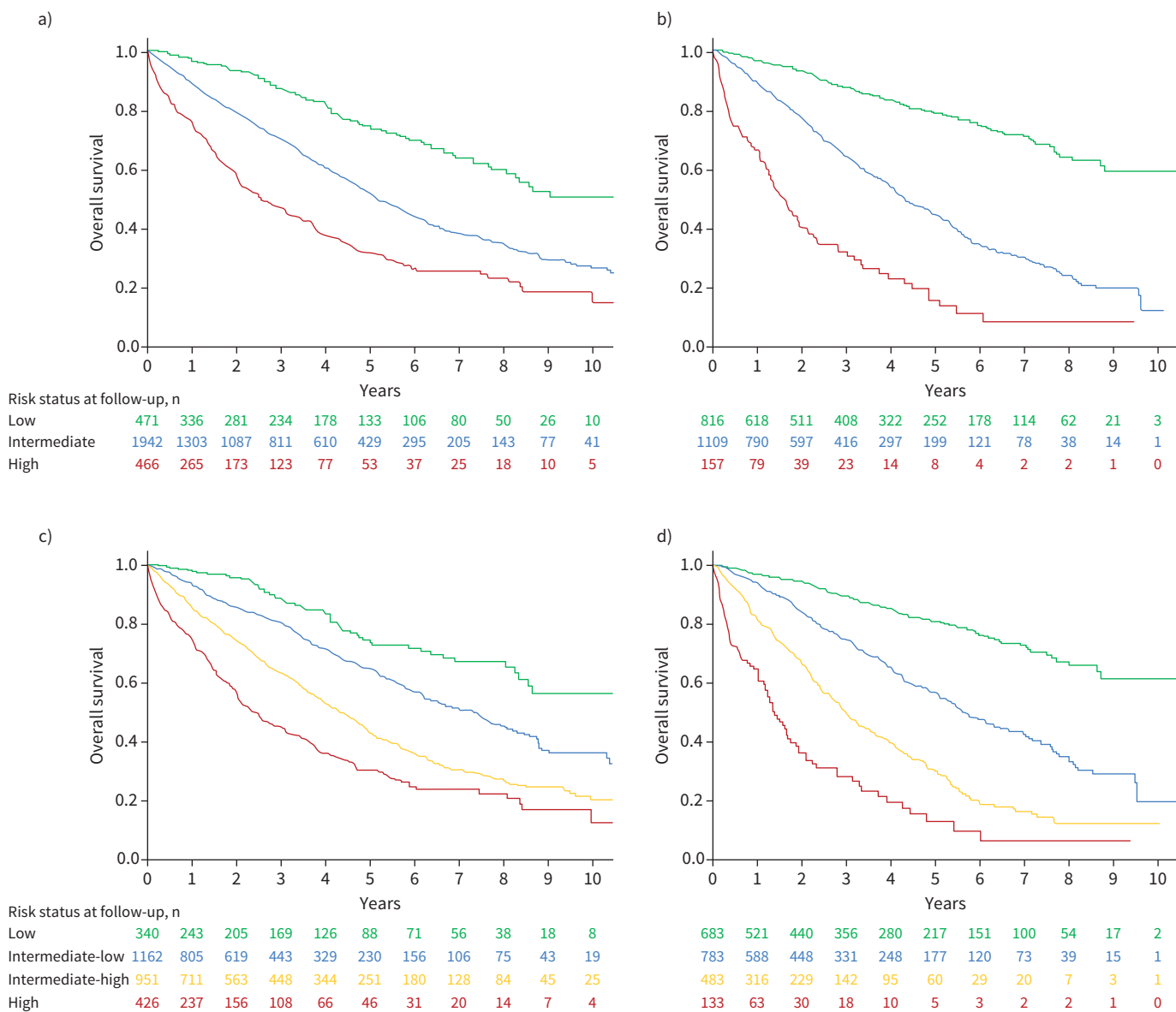


FIGURE 2 Survival according to a three-stratum strategy a) at diagnosis and b) after first reassessment. Survival according to the four-stratum risk assessment strategy c) after diagnosis and d) after first reassessment. Log rank test $p < 0.001$ for all models.

10% experienced early mortality or underwent lung transplantation before a full reassessment and 18% had no reassessment of risk available. 29% of patients changed risk categories (by improving or worsening) between baseline and follow-up, with 10% remaining as stable low risk, 31% as stable intermediate risk and 3% remaining as high risk. Few high risk patients improved to the low risk category.

Using the four-stratum method, the proportion of patients classified as low risk also improved from baseline (12%) to follow-up (24%). A higher proportion of patients changed risk category at follow-up when using the four-stratum method (39%) than with the three-stratum method (29%). 10% of patients worsened by at least one category; 32% improved by at least one category; 10% were stable in the low risk category; and 3% of patients were “stable” in the high risk category. Of the intermediate-low risk patients at baseline, 39% changed risk categories, 39% stayed at intermediate-low risk, 6% experienced early death or transplant and 15% had no follow-up available. Of the patients who were at intermediate-high risk at baseline, ~48% changed risk categories, 25% remained intermediate-high risk, 10% experienced early death or transplant and 18% had no follow-up available. The proportion of patients who were high risk and

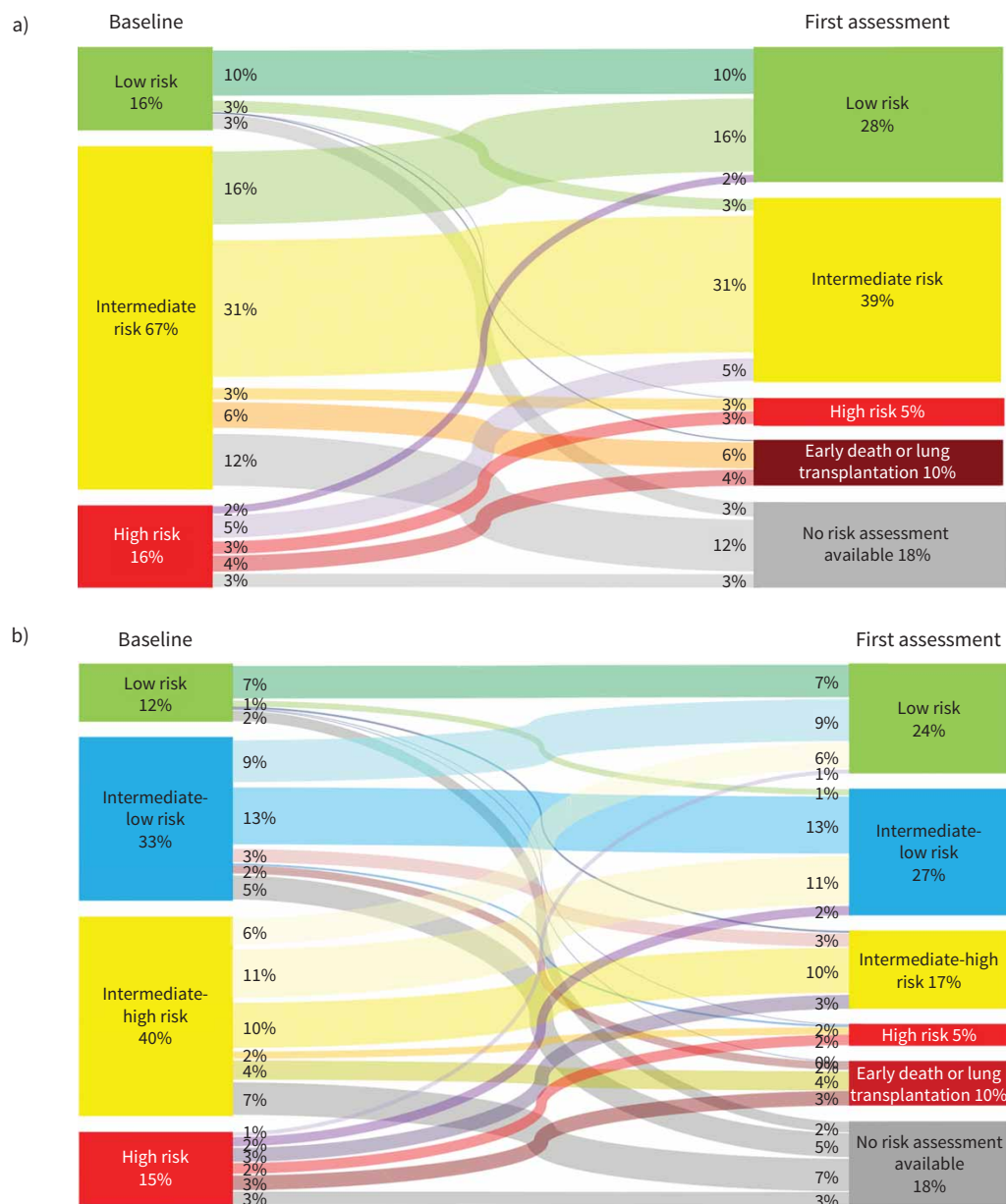


FIGURE 3 Sankey diagrams showing changes in risk status using the a) three-stratum method and b) four-stratum method. Sankey diagrams are a visualisation technique to display flows. Each panel shows the flow of patients between risk strata (nodes) from baseline to first reassessment. The width of each band is weighted to the proportion of patients who had a given risk trajectory.

remained high risk was similar using the three- or four-stratum approach. Among patients who were high risk at baseline, 20% improved to low risk or intermediate-low risk.

Survival was similar between patients who were stable in the low risk category and those who improved to low risk, whereas there were clear differences in long-term survival between those who ended up at intermediate-low risk compared to those who were intermediate-high risk at follow-up (figure 4). Compared to patients who remained stable in the intermediate-low risk category or worsened from low to intermediate-low, survival was incrementally worse for those who improved to intermediate-low from higher risk groups and for those who improved to intermediate-high risk. Persistent high risk status or worsening to high risk was associated with the worst outcome. Survival according to the evolution in risk category is also shown for each baseline risk group in figure 5.

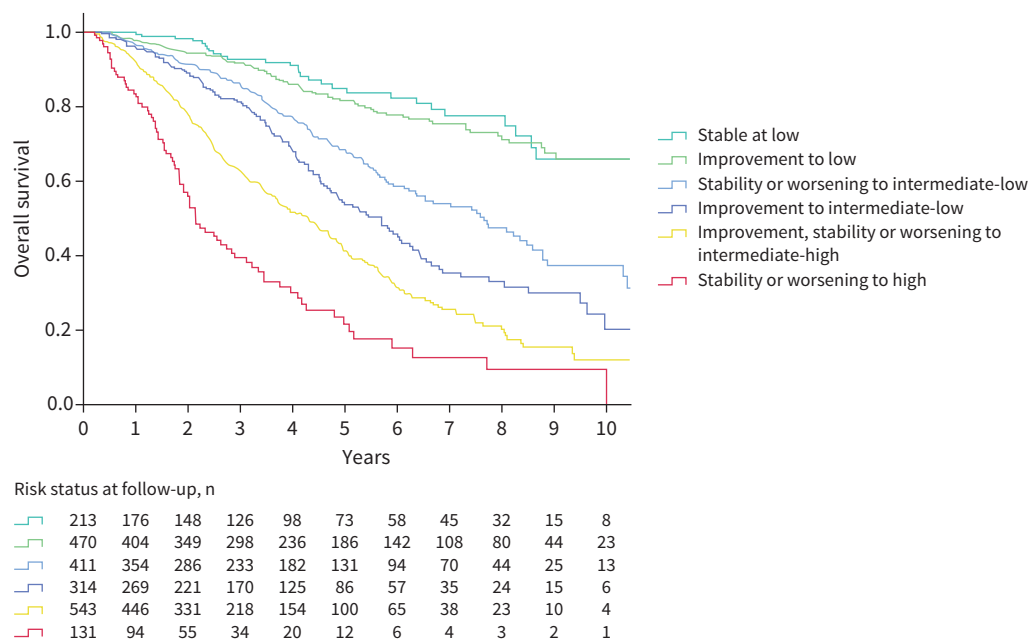


FIGURE 4 Overall survival according to changes in risk strata between baseline and first reassessment. Log rank test $p < 0.001$.

Four-stratum risk assessment in PAH subgroups

Survival at baseline and follow-up according to the four-stratum method for the subgroups with idiopathic/heritable/drug- and toxin-induced PAH, CTD-PAH and systemic sclerosis (SSc)-associated PAH are presented in supplementary figure E2, and for portopulmonary hypertension (PoPH) in supplementary figure E3. There were significant differences across risk groups using the four-stratum method for all subgroup populations (log rank test $p < 0.01$ for all comparisons). In addition, we found an increase in the proportion of patients treated with initial dual combination therapy and fewer treated with monotherapy in the patients diagnosed in 2015 or later (supplementary table E2). Risk stratification models with three and four strata performed similarly in the pre-2015 and 2015–2020 groups (supplementary table E3). A sensitivity analysis excluding the 295 patients who died within 1 year of diagnosis did not change the overall results (supplementary table E4).

Discussion

In this large cohort of incident PAH patients from the French PH registry, we evaluated a refined four-stratum risk assessment approach, based on new cut-points for 6MWD and NT-proBNP/BNP, and compared this to a three-stratum risk assessment method previously proposed by the SPHAR/COMPERA investigators. Our main findings were that 1) few patients were low risk at baseline with either approach, and $< 40\%$ achieved the treatment goal of a low risk profile during follow-up, regardless of which method was used; 2) using a four-stratum model identified distinct groups within the intermediate risk category, with an intermediate-low risk group that had $< 10\%$ 1-year mortality and an intermediate-high risk group with a $> 10\%$ 1-year mortality risk; 3) the four-stratum risk model had modestly higher discrimination for long-term mortality and 1-year mortality compared to the three-stratum model; 4) a greater proportion of patients changed risk category between baseline and follow-up when using the four-stratum approach compared to the three-stratum approach; 5) changes in risk category were associated with survival, with a more nuanced assessment of survival possible according to permutations of changes in the four-stratum risk method; and 6) there were differences in survival across the four risk strata in all subgroups of patients with PAH. Our study confirms a recent analysis by the COMPERA investigators and provides new additional analyses to support the statistical validity of this approach. The four-stratum method was more sensitive in assessing changes in risk after initial treatment and was superior at discriminating long-term and short-term (1-year) mortality, which will help patients and clinicians make better informed decisions about treatment.

Achieving or maintaining a low risk profile is the therapeutic objective for patients with PAH [1, 17, 18]. Risk prediction is essential to inform patients about their prognosis and guides clinical decision making [1, 17]. There are several useful PAH risk assessment tools available, each with advantages and disadvantages.

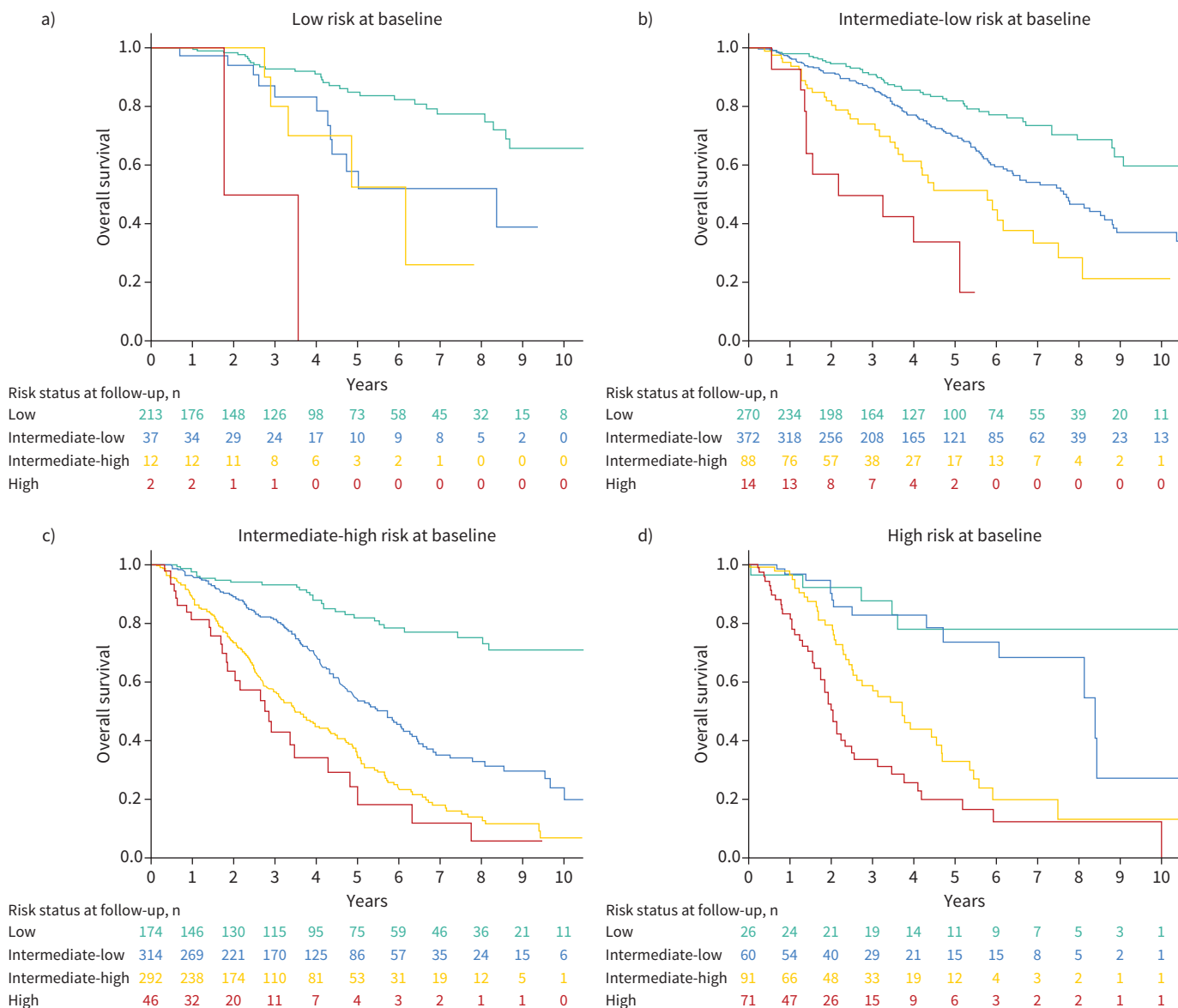


FIGURE 5 Survival according to change in risk strata for patients who were a) low risk at baseline, b) intermediate-low risk at baseline, c) intermediate-high risk at baseline and d) high risk at baseline. Log-rank test $p < 0.001$ for each panel.

Importantly, objective multivariable risk scores are better at predicting a patient’s risk than clinical gestalt, which is why such tools are essential in modern clinical practice [19, 20]. We found that by using a four-stratum approach with three variables, as proposed by the COMPERA and SPAHR investigators [14, 15], a greater proportion of patients changed their risk classification, although this alone does not necessarily indicate a better tool. The more important observation is that risk assessment using the four-stratum classification identified groups of patients in intermediate-low and intermediate-high risk categories who had clearly different outcomes. Also, the four-stratum method seemed to have higher discrimination of short and long-term outcomes compared to the three-stratum SPAHR/COMPERA approach. This indicates that a more nuanced categorisation of risk can refine prediction of long-term survival using the four-stratum method. Our findings also highlight the importance of achieving a low risk status regardless of where a patient starts, but also the prognostic relevance of unsatisfactory treatment responses and worsening risk status despite initial treatment.

Our cohort was larger than the recent COMPERA 2.0 cohort which derived the four-stratum approach, spanned a similar contemporary time period, and was comparable in terms of patient characteristics and haemodynamic severity. In validating the COMPERA four-stratum method, we confirm its simplicity and

its utility in identifying a greater proportion of patients who changed risk over time. Our study also builds upon the COMPERA 2.0 analysis, with our statistical modelling demonstrating good discrimination for short and long-term survival using the four-stratum approach. The four-stratum method overlaps considerably with REVEAL 2.0 Lite, and indeed is based on the 6MWD and NT-proBNP cut-points proposed in REVEAL 2.0 Lite [11]. The COMPERA 2.0 approach uses three variables, whereas REVEAL 2.0 Lite uses six variables (WHO FC, 6MWD, NT-proBNP/BNP, systolic blood pressure, heart rate and renal function), and there is more granularity with four NT-proBNP groups with the COMPERA 2.0 four-stratum approach as opposed to three NT-proBNP groups in REVEAL 2.0 Lite. These new cut-points for NT-proBNP were data-derived from the COMPERA 2.0 derivation study. In our study, the discrimination of the four-stratum model at follow-up was comparable to that reported for the six-variable REVEAL 2.0 Lite score overall (C-index 0.73) and when those variables that differ from the COMPERA 2.0 score (systolic blood pressure, heart rate and renal function) were missing (C-index 0.72) [11]. Thus, our data indirectly support the REVEAL 2.0 Lite model and confirm the validity of the COMPERA 2.0 approach. Regardless of which method is used, a key message from all investigators and guidelines is to perform risk assessment on a regular and recurring basis.

One criticism of the three-stratum SPAHR/COMPERA risk assessment approach is that most patients remain at intermediate risk, which is potentially problematic in clinical practice. The treatment algorithm proposed at the 6th World Symposium on Pulmonary Hypertension recommends treatment escalation, which includes parenteral prostacyclin analogues or lung transplantation assessment for patients who remain at intermediate risk despite optimal therapy [12]. The management of low risk patients and high risk patients is relatively straightforward, since low risk patients are achieving treatment goals and high risk patients clearly require more aggressive interventions such as parenteral therapies and/or lung transplantation referral, or should be provided with palliative care options if they are not candidates for these interventions. Refining risk within the intermediate risk category using the four-stratum approach will be valuable to clinicians and will better inform treatment decisions for this group, especially with respect to transplantation or parenteral prostanoids. For example, lung transplantation in general should be considered for patients with an estimated $\geq 50\%$ mortality at 2 years [21]. Using the three-stratum method, the median survival time for patients who remained at intermediate risk at follow-up in our cohort was 4.3 years (IQR 2.1–7.8 years), so referral for transplantation would be premature for most of these intermediate risk patients. Using the four-stratum method at follow-up, median survival was 5.8 years (IQR 3.0–9.6 years) for intermediate-low risk patients and 3 years (IQR 1.4–5.5 years) for intermediate-high risk patients. Therefore, lung transplantation assessment would certainly be reasonable for patients who remained at intermediate-high risk after initial therapy, but likely not reasonable for those intermediate-low risk patients. Another scenario in which the distinction between intermediate-low and intermediate-high may be clinically useful pertains to the addition of a third medication to a patient already on dual oral therapy with a phosphodiesterase type-5 inhibitor and endothelin receptor antagonist. In places where oral selexipag or oral treprostinil are available, it might be more acceptable to add an oral third agent for an intermediate-low risk patient, whereas a more compelling case can be made for adding a parenteral prostanoid for intermediate-high risk patients. These examples illustrate the value of refining risk within the intermediate risk group of patients, with respect to clinical decision making.

A strength of this study was the large cohort size, which permitted validation of the risk assessment methods in several important subgroups such as idiopathic, heritable, drug- and toxin-induced, CTD-associated, SSc-associated and congenital heart disease associated PAH, which often differ in terms of characteristics, treatment and prognosis. The limitations of this study include the retrospective nature of the analysis and missing data for follow-up risk assessment for 28% of patients, due to early death or transplantation (10%) or lack of available follow-up data (18%). There may also have been changes in medical therapies after the first follow-up assessment, and the impact of subsequent therapeutic decisions on long-term risk and long-term survival were not accounted for in this analysis, which is a limitation. While the four-stratum model performed well at discriminating outcomes, C-statistics in the range of 0.6–0.7 are considered good, but not excellent. As most PAH risk scores have C-statistics in the range of 0.6–0.8 [8, 12, 13], future studies should aim to improve the performance of risk prediction methods, ideally without sacrificing simplicity. We also noted that the four-stratum approach may be less useful in the subgroup with PoPH, with little difference between the intermediate-high and high risk strata (supplementary figure E3). The optimal stratification method for patients with PoPH requires further study, as survival in this population is highly dependent on other factors, such as the presence of cirrhosis and severity of liver disease [22, 23]. Lastly, we included only patients with mPAP ≥ 25 mmHg, which was the accepted threshold for PAH at the time of this cohort. The applicability of our results to populations with PAH defined as mPAP > 20 mmHg according to the 6th World Symposium on Pulmonary Hypertension requires further study [24].

In conclusion, our study supports the notion of a four-stratum risk assessment approach using revised cut-points for the 6MWD and NT-proBNP/BNP. The four-stratum approach better discriminated the risk of future mortality over the short and long term and appears to have greater sensitivity in identifying changes in risk. This approach enhanced the granularity of risk assessment, especially for intermediate risk patients, which will likely help clinicians evaluate more subtle treatment-related improvements and better identify patients who require more aggressive treatment. Further work is needed to determine which risk assessment method is most sensitive to change in the context of clinical trials.

Acknowledgements: We wish to acknowledge Laurence Rottat (Hôpital Bicêtre, Le Kremlin Bicêtre, France) for her help in obtaining the data for this study and her work in managing data in the French PH registry. We wish to acknowledge all members from the French Reference Center for Pulmonary Hypertension (PulmoTension) and from the French Reference Network for Rare Respiratory Diseases (RespiFIL). We especially thank Marius Hoepfer (Hannover Medical School, Hannover, Germany), Stephan Rozenkrantz (University of Cologne, Cologne, Germany), and the COMPERA investigators for the fruitful discussions and collaboration leading to this manuscript.

Conflict of interest: A. Boucly reports personal fees from Actelion, Bayer and Merck, outside the submitted work. J. Weatherald reports grants, personal fees and non-financial support from Janssen Inc., grants, personal fees and non-financial support from Actelion, personal fees and non-financial support from Bayer, personal fees from Novartis, outside the submitted work. L. Savale reports personal fees from Actelion, personal fees from MSD, grants and personal fees from GSK, outside the submitted work. P. de Groote reports consulting fees from Actelion, Janssen, MSD, Novartis, Servier, Boehringer Ingelheim, Abbott, Boston, AstraZeneca, Bayer; lecture honoraria from Abbott, Vifor, MSD, Servier, Novartis, AstraZeneca, Actelion, Janssen, Medtronic; outside the submitted work. V. Cottin reports advisory board fees and non-financial support from Actelion, advisory board fees from Bayer/MSD, outside the submitted work. G. Prévot reports personal fees from Actelion and GSK, outside the submitted work. A. Chaouat reports consulting fees from GSK, Actelion and Bayer, outside the submitted work. F. Picard has nothing to disclose. D. Horeau-Langlard reports grants from Acceleron, outside the submitted work. A. Bourdin reports grants from AstraZeneca and Boehringer Ingelheim; consulting fees from AstraZeneca, GSK, Novartis, Boehringer Ingelheim, Chiesi, Sanofi Regeneron, Amgen; lecture honoraria from AstraZeneca, GSK, Novartis, Boehringer Ingelheim, Chiesi, Sanofi Regeneron, Roche; travel support from Boehringer Ingelheim, Chiesi, Sanofi Regeneron, AstraZeneca, GSK, Roche; participation on advisory boards at AB science, AstraZeneca, GSK, Sanofi Regeneron, Novartis, Acceleron; and acted as investigator in clinical trials for Vertex, Abbvie, Galapagos, Fibrogen, Nuaira, PulmonX, Gossamer, Acceleron; outside the submitted work. E-M. Jutant has nothing to disclose. A. Beurnier has nothing to disclose. M. Jevnikar has nothing to disclose. X. Jais reports grants from Bayer, Janssen and Merck; lecture honoraria from Janssen and Merck; outside the submitted work. G. Simonneau reports grants and personal fees from Janssen (formerly Actelion), Bayer and MSD; personal fees from Acceleron, outside the submitted work. D. Montani reports grants from Acceleron, Janssen and Merck; steering committee fees from Acceleron; lecture honoraria from Bayer, Janssen and Merck; outside the submitted work. O. Sitbon reports grants from Acceleron, Janssen, GSK and MSD; steering committee fees from Gossamer Bio, Janssen and MSD; lecture honoraria from AOP Orphan, Janssen, Ferrer and MSD; advisory board participation at Acceleron, Janssen and MSD; outside the submitted work. M. Humbert reports grants, steering committee consulting fees, and advisory board participation from Acceleron, Janssen and Merck; lecture honoraria from AOP, Janssen and Merck; steering committee participation at United Therapeutics; outside the submitted work.

References

- 1 Galiè N, Humbert M, Vachiery J-L, *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). *Eur Respir J* 2015; 46: 903–975.
- 2 Hoepfer 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.
- 3 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.
- 4 Boucly A, Weatherald J, Savale L, *et al.* Risk assessment, prognosis and guideline implementation in pulmonary arterial hypertension. *Eur Respir J* 2017; 50: 1700889.
- 5 Weatherald J, Boucly A, Launay D, *et al.* Haemodynamics and serial risk assessment in systemic sclerosis associated pulmonary arterial hypertension. *Eur Respir J* 2018; 52: 1800678.
- 6 Hoepfer MM, Pittrow D, Opitz C, *et al.* Risk assessment in pulmonary arterial hypertension. *Eur Respir J* 2018; 51: 1702606.
- 7 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.

- 8 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.
- 9 Humbert M, Farber HW, Ghofrani H-A, *et al.* Risk assessment in pulmonary arterial hypertension and chronic thromboembolic pulmonary hypertension. *Eur Respir J* 2019; 53: 1802004.
- 10 Benza RL, Farber HW, Frost AE, *et al.* Application of the REVEAL risk score calculator 2.0 in the PATENT study. *Int J Cardiol* 2021; 332: 189–192.
- 11 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.
- 12 Boucly A, Weatherald J, Humbert M, *et al.* Risk assessment in pulmonary arterial hypertension. *Eur Respir J* 2018; 51: 1800279.
- 13 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.
- 14 Kylhammar D, Hjalmarsson C, Hesselstrand R, *et al.* Predicting mortality during long-term follow-up in pulmonary arterial hypertension. *ERJ Open Res* 2021; 7: 00837-2020.
- 15 Hoeper MM, Pausch C, Olsson KM, *et al.* COMPERA 2.0: a refined 4-stratum risk assessment model for pulmonary arterial hypertension. *Eur Respir J* 2021; in press [<https://doi.org/10.1183/13993003.02311-2021>].
- 16 Newson RB. Comparing the predictive powers of survival models using Harrell's C or Somers' D. *Stata J* 2010; 10: 339–358.
- 17 Galiè N, Channick RN, Frantz RP, *et al.* Risk stratification and medical therapy of pulmonary arterial hypertension. *Eur Respir J* 2019; 53: 1801889.
- 18 Weatherald J, Boucly A, Sahay S, *et al.* The low-risk profile in pulmonary arterial hypertension. Time for a paradigm shift to goal-oriented clinical trial endpoints? *Am J Respir Crit Care Med* 2018; 197: 860–868.
- 19 Simons JE, Mann EB, Pierozynski A. Assessment of risk of disease progression in pulmonary arterial hypertension: insights from an international survey of clinical practice. *Adv Ther* 2019; 36: 2351–2363.
- 20 Sahay S, Tonelli AR, Selej M, *et al.* Risk assessment in patients with functional class II pulmonary arterial hypertension: comparison of physician gestalt with ESC/ERS and the REVEAL 2.0 risk score. *PLoS One* 2020; 15: e0241504.
- 21 Leard LE, Holm AM, Valapour M, *et al.* Consensus document for the selection of lung transplant candidates: an update from the International Society for Heart and Lung Transplantation. *J Heart Lung Transplant* 2021; 40: 1359–1379.
- 22 Savale L, Guimas M, Ebstein N, *et al.* Portopulmonary hypertension in the current era of pulmonary hypertension management. *J Hepatol* 2020; 73: 130–139.
- 23 Le Pavec J, Souza R, Herve P, *et al.* Portopulmonary hypertension: survival and prognostic factors. *Am J Respir Crit Care Med* 2008; 178: 637–643.
- 24 Simonneau G, Montani D, Celermajer DS, *et al.* Haemodynamic definitions and updated clinical classification of pulmonary hypertension. *Eur Respir J* 2019; 53: 1801913.