

# Risk stratification in pulmonary arterial hypertension using Bayesian analysis

Manreet K. Kanwar <sup>1</sup>, Mardi Gomberg-Maitland<sup>2</sup>, Marius Hoepfer <sup>3</sup>, Christine Pausch<sup>4</sup>, David Pittrow <sup>5</sup>, Geoff Strange<sup>6</sup>, James J. Anderson<sup>7</sup>, Carol Zhao<sup>8</sup>, Jacqueline V. Scott<sup>9</sup>, Marek J. Druzdzal<sup>10</sup>, Jidapa Kraisingka<sup>11</sup>, Lisa Lohmueller<sup>12</sup>, James Antaki<sup>13</sup> and Raymond L. Benza<sup>14</sup>

**Affiliations:** <sup>1</sup>Cardiovascular Institute at Allegheny Health Network, Pittsburgh, PA, USA. <sup>2</sup>George Washington School of Medicine and Health Sciences, Washington, DC, USA. <sup>3</sup>Dept of Respiratory Medicine, Hannover Medical School, German Center for Lung Research (DZL), Hannover, Germany. <sup>4</sup>GWT-TUD GmbH, Dresden, Germany. <sup>5</sup>Faculty of Institute for Clinical Pharmacology, Technical University, Dresden, Germany. <sup>6</sup>School of Medicine, University of Notre Dame, Fremantle, Australia. <sup>7</sup>Respiratory Dept, Sunshine Coast University Hospital, Nambour, Australia. <sup>8</sup>Actelion Pharmaceuticals US, A Janssen Pharmaceutical Company of Johnson & Johnson, San Francisco, CA, USA. <sup>9</sup>School of Biomedical Engineering, Carnegie Mellon University, Pittsburgh, PA, USA. <sup>10</sup>Faculty of Computer Science, Bialystok University of Technology, Bialystok, Poland. <sup>11</sup>Faculty of Information and Communication Technology, Mahidol University, Salaya, Thailand. <sup>12</sup>Language Technologies Institute, Carnegie Mellon University, Pittsburgh, PA, USA. <sup>13</sup>Meinig School of Biomedical Engineering, Cornell University, Ithaca, NY, USA. <sup>14</sup>Ohio State Medical Center, Columbus, OH, USA.

**Correspondence:** Corresponding author: Raymond L. Benza, 320 E North Avenue, Pittsburgh, Pennsylvania – 15212, USA. Email: raymond.benza@osumc.edu



@ERSpublications

**Bayesian machine-learning algorithms can improve discrimination of risk stratification in PAH. Our BN model, titled PHORA, predicts 1-year mortality with an AUC of 0.8, risk-stratifies patients effectively and is validated in two independent PAH registries.** <https://bit.ly/2xc0EVJ>

**Cite this article as:** Kanwar MK, Gomberg-Maitland M, Hoepfer M, *et al.* Risk stratification in pulmonary arterial hypertension using Bayesian analysis. *Eur Respir J* 2020; 56: 2000008 [<https://doi.org/10.1183/13993003.00008-2020>].

## ABSTRACT

**Background:** Current risk stratification tools in pulmonary arterial hypertension (PAH) are limited in their discriminatory abilities, partly due to the assumption that prognostic clinical variables have an independent and linear relationship to clinical outcomes. We sought to demonstrate the utility of Bayesian network-based machine learning in enhancing the predictive ability of an existing state-of-the-art risk stratification tool, REVEAL 2.0.

**Methods:** We derived a tree-augmented naïve Bayes model (titled PHORA) to predict 1-year survival in PAH patients included in the REVEAL registry, using the same variables and cut-points found in REVEAL 2.0. PHORA models were validated internally (within the REVEAL registry) and externally (in the COMPERA and PHSANZ registries). Patients were classified as low-, intermediate- and high-risk (<5%, 5–20% and >10% 12-month mortality, respectively) based on the 2015 European Society of Cardiology/European Respiratory Society guidelines.

**Results:** PHORA had an area under the curve (AUC) of 0.80 for predicting 1-year survival, which was an improvement over REVEAL 2.0 (AUC 0.76). When validated in the COMPERA and PHSANZ registries, PHORA demonstrated an AUC of 0.74 and 0.80, respectively. 1-year survival rates predicted by PHORA were greater for patients with lower risk scores and poorer for those with higher risk scores ( $p < 0.001$ ), with excellent separation between low-, intermediate- and high-risk groups in all three registries.

**Conclusion:** Our Bayesian network-derived risk prediction model, PHORA, demonstrated an improvement in discrimination over existing models. This is reflective of the ability of Bayesian network-based models to account for the interrelationships between clinical variables on outcome, and tolerance to missing data elements when calculating predictions.

Received: 9 Jan 2020 | Accepted after revision: 22 April 2020

Copyright ©ERS 2020

## Introduction

Pulmonary arterial hypertension (PAH) is a rapidly progressive, incurable disease with a median survival of ~7 years after diagnosis [1]. Accurate risk stratification in PAH accommodates demographic, clinical, haemodynamic and functional parameters, allowing clinicians to identify treatment goals, monitor disease progression and facilitate timely referral to a PAH centre and/or lung transplantation [2]. Large PAH patient registries in Europe and United States have been used to develop PAH risk scores to quantify these predictions [2, 3]. These include algorithms derived from the 2015 European Society of Cardiology (ESC)/European Respiratory Society (ERS) guidelines using derivation cohorts from the French Pulmonary Hypertension Registry, Swedish PAH Register and the Comparative, Prospective Registry of Newly Initiated Therapies for Pulmonary Hypertension (COMPERA); as well as the United States Registry to Evaluate Early and Long-Term PAH Disease Management (REVEAL) risk equation and calculator [3–7]. Although derived from contemporary patient registries, their associated discriminatory abilities, are fair to good at best, limiting their overall use in clinical practice. One of the important limitations of existing risk-stratification tools includes the assumption that pertinent and prognostic clinical variables have an independent and linear relationship to a particular outcome measure, without intervariable relationships.

Bayesian networks are highly efficient and sophisticated algorithms derived using data mining (a process of discovering patterns in pre-existing data). Such networks can be trained to recognise complex medical data in a time-efficient manner, thereby acting as a tool for predicting clinical outcomes based on learned information. They can account for dynamic, nonlinear interactions between multiple variables and their interdependency in influencing outcomes at various time points. These networks can encode both qualitative and quantitative knowledge, can be represented diagrammatically or numerically and provide a rigorous framework to perform inferences from predictive variables [8]. In this article, we sought to demonstrate the utility of Bayesian network-based machine learning in enhancing the predictive ability of a contemporary risk stratification tool, REVEAL 2.0 [9].

## Methods

### *Patient population/derivation cohort*

The REVEAL registry design and development of risk calculators have been described previously [10]. In brief, the observational, prospective REVEAL registry included PAH patients from 55 hospitals based in the United States. Patients analysed in the REVEAL registry include 74% previously diagnosed and 26% newly diagnosed PAH patients. REVEAL was conducted in accordance with the amended Declaration of Helsinki. Institutional review boards at each study site approved the protocol and written, informed consent was obtained from all patients. Our Bayesian network models were derived from the final study data of REVEAL 2.0 [9] and included PAH patients who survived  $\geq 1$ -year post-enrolment to allow sufficient capture of all-cause hospitalisation data in the previous 6 months (derivation cohort).

### *Model development with Bayesian networks*

Bayesian networks incorporate relationships and processes in individual patient data within a large dataset to predict probability of the outcomes for survival and adverse events. For our analysis, we used tree-augmented naïve (TAN) Bayes algorithms for structure and parameter learning [8, 11]. TAN architecture adds a level of complexity to the simplest network form (a naïve Bayes), allowing independent variables to both directly and indirectly impact the outcome through their influence on other variables. These inferences are represented diagrammatically, in which nodes represent pertinent variables and directed arrows between nodes represent interactions between those variables. Absence of an arrow between a pair of nodes implies independence between those variables. Only patients who had data at the 1-year mark available were included, using variables at 12 months, if available. If there was no assessment done at 1 year, the variable most recent to that time point (including assessment at enrolment, up to 12 months) was used. Our TAN model was structured from the same database, variables and cut-points found in the REVEAL 2.0 calculator, looking at survival at 12 months as the clinical outcome (table 1). Clinical variables were coded as nodes, which were then discretised into prespecified intervals (e.g. N-terminal pro-brain natriuretic peptide levels (<300, 300–1100, >1100 pg·mL<sup>-1</sup>) or 6-min walk distance (<165, 165–320, 320–440, >440 m)), as required for Bayesian methodology. The Bayesian network model learned the direction and magnitude of influence between these prespecified variables on each other as well as the final clinical outcome, represented in the model as conditional probability tables. The final model represents the joint probability distribution over its variables, by taking the product of all prior and conditional probability distributions (figure 1). We named the derived model the Pulmonary Hypertension Outcomes Risk Assessment (PHORA). We created all the models described in this paper using GeNIe software developed at the University of Pittsburgh. GeNIe is a machine-learning software which provides a platform for artificial intelligence modelling based on Bayesian networks ([www.bayesfusion.com/](http://www.bayesfusion.com/)).

TABLE 1 List of variables and their discrete states from the Registry to Evaluate Early and Long-Term PAH Disease Management (REVEAL 2.0) risk score

Risk factor in REVEAL	Random variable	Nodes in Bayesian network
CTD-PAH Heritable POPH	WHO group I	CTD Heritable POPH Other
Male >60 years	Sex Age years	Female Male ≤60 >60
Comorbidity eGFR <60 mL·min <sup>-1</sup> ·1.73 m <sup>-2</sup> or renal insufficiency if eGFR is unavailable	eGFR <60 mL·min <sup>-1</sup> ·1.73 m <sup>-2</sup> or renal insufficiency if eGFR is unavailable	Yes No
NYHA functional class I NYHA functional class III NYHA functional class IV	NYHA/WHO functional class	I II III IV
Systolic blood pressure <110 mmHg	Systolic blood pressure mmHg	<110 ≥110
Heart rate >96 bpm	Heart rate bpm	≤96 >96
6MWD <165 m 6MWD 320–440 m 6MWD ≥440 m	6MWD m	<165 165 to <320 320–440 ≥440
BNP <50 pg·mL <sup>-1</sup> or NT-proBNP <300 pg·mL <sup>-1</sup> BNP 200 to <800 pg·mL <sup>-1</sup> BNP ≥800 pg·mL <sup>-1</sup> or NT-proBNP ≥1100 pg·mL <sup>-1</sup>	BNP or NT-proBNP pg·mL <sup>-1</sup>	<50 or <300 50–200 or 300–1100 200–800 ≥800 or ≥1100
Echocardiogram Pericardial effusion	Pericardial effusion	Yes No
Right heart catheterisation Mean RAP ≥20 mmHg within 1 year	Mean RAP	<20 ≥20
PVR <5 WU	PVR WU	<5 ≥5
PFT <i>D</i> <sub>LCO</sub> <40%	% <i>D</i> <sub>LCO</sub>	<40 ≥40
Hospitalisation in 6 months	Hospitalisation in 6 months	Yes No
Survival at 12 months	Survival at 12 months	Yes No

CTD-PAH: connective tissue disease-associated pulmonary arterial hypertension; POPH: portopulmonary hypertension; eGFR: estimated glomerular filtration rate; NYHA: New York Heart Association; 6MWD: 6-min walk distance; BNP: brain natriuretic peptide; NT-proBNP: N-terminal pro-BNP; RAP: right atrial pressure; PVR: pulmonary vascular resistance; WU: Wood Units; PFT: pulmonary function test; *D*<sub>LCO</sub>: diffusing capacity of the lungs for carbon monoxide; WHO: World Health Organization.

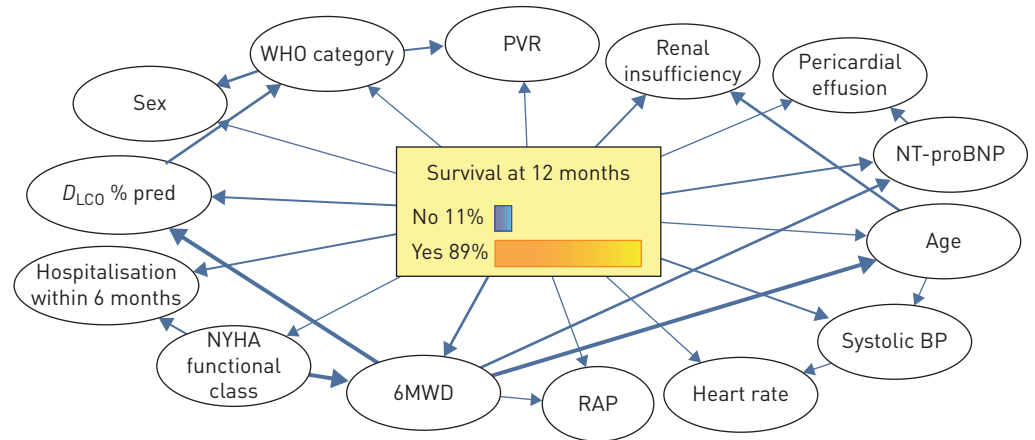


FIGURE 1 Structure of the Pulmonary Hypertension Outcomes Risk Assessment (PHORA) Bayesian network model, with conditional probability table for survival. PVR: pulmonary vascular resistance; NT-proBNP: N-terminal pro-BNP; BP: blood pressure; RAP: right atrial pressure; 6MWD: 6-min walk distance; NYHA: New York Heart Association;  $D_{LCO}$ : diffusing capacity of the lungs for carbon monoxide; WHO: World Health Organization.

### Patient population/validation cohorts

We validated the PHORA Bayesian network model both internally and externally, utilising the following cohorts and methodologies. Internal validation: we validated the PHORA model internally within the REVEAL registry using 10-fold cross-validation and report the results of this validation as AUC. External validation: we validated the PHORA model externally in two registries: 1) the COMPERA registry, which is an ongoing multinational European registry comprised of patients with pulmonary hypertension/PAH enrolled since May 2007 [4]. The PHORA model was validated on 3849 newly diagnosed, consecutively enrolled PAH patients. Data from time of enrolment were considered; 2) the Pulmonary Hypertension Society of Australia and New Zealand (PHSANZ) Registry, which collects data from patients with all subgroups of pulmonary hypertension since December 2011 from 16 Australian and two New Zealand centres [12]. PHORA was validated in those PAH patients who had 1-year data available (978 out of 1076). Variables included were at the time closest to 1-year mark, as available (similar to REVEAL 2.0 and PHORA). These included both previously (75%) and newly diagnosed (25%) PAH patients within the PHSANZ registry.

PHORA performance in predicting survival in each registry was measured using the AUC method. Kaplan–Meier curves were then derived for the PHORA-predicted mortality risk thresholds (*i.e.* low risk <5% 12-month mortality; intermediate risk 5–10% 12-month mortality; high risk >10% 12-month mortality) based on the 2015 ESC/ERS guidelines [4]. The statistical significance of the ability of PHORA to stratify risk groups in each of the three registry populations was calculated using Chi-squared analysis (SPSS; IBM, Armonk, NY, USA).

### Results

Of the 3515 patients enrolled in REVEAL, 2529 were in the registry at 12 months after enrolment and included in the PHORA derivation model. Of these, 73.7% were previously diagnosed (*i.e.* >3 months before enrolment) and 26.3% were newly diagnosed (*i.e.* ≤3 months before enrolment). The majority of the patients were female (80%), New York Heart Association/World Health Organization functional class II (41.3%) or III (45.9%), with a mean age of 53.6 years. The clinical variables across all three registries (REVEAL, COMPERA and PHSANZ) are presented in table 2.

Revising REVEAL 2.0 to a Bayesian network (PHORA; figure 1) improved the predictive power of the calculator. The AUC of 0.80 for predicting 1-year survival for PHORA indicated improved discrimination in predicting mortality over REVEAL 2.0 (0.76, 95% CI 0.74–0.78) and REVEAL 1.0 (0.71, 95% CI 0.68–0.77). PHORA had specificity of 0.76 (95% CI 0.69–0.84), sensitivity of 0.79 (95% CI 0.72–0.82), negative predictive value of 0.30 (95% CI 0.25–0.34) and positive predictive value of 0.97 (95% CI 0.96–0.98) for 1-year survival. PHORA demonstrated an AUC of 0.74 and 0.80 when validated in the COMPERA and PHSANZ registries, respectively (figure 2). Hence, PHORA outperformed the contemporary REVEAL 2.0 risk stratification model.

Patients were classified as low risk (<5% 12-month mortality); intermediate risk (5–10% 12-month mortality) and high risk (>10% 12-month mortality) based on the 2015 ESC/ERS guidelines. 12-month

TABLE 2 Clinical variables through three pulmonary arterial hypertension registries: Registry to Evaluate Early and Long-Term PAH Disease Management (REVEAL), Comparative Prospective Registry of Newly Initiated Therapies for Pulmonary Hypertension (COMPERA) and Pulmonary Hypertension Society of Australia and New Zealand (PHSANZ) registry

	REVEAL	COMPERA	PHSANZ
<b>Patients n</b>	2529	3849	978
<b>Sex</b>			
Male	505 (20.0)	1373 (35.7)	218 (22.3)
Female	2024 (80.0)	2476 (64.3)	760 (77.7)
<b>Age years</b>			
<60	1673 (66.1)	1661 (43.2)	392 (40.1)
>60	856 (33.9)	2188 (56.8)	586 (59.9)
<b>WHO category</b>			
Idiopathic	1171 (46.3)	2158 (46.3)	468 (47.9)
CTD-PAH	649 (25.7)	1347 (35.0)	315 (32.2)
POPH-PAH	139 (5.5)	188 (4.9)	26 (2.7)
FPAH	74 (2.9)	76 (2.0)	20 (2.0)
Other	496 (7.9)	80 (2.1)	144 (14.7)
<b>NYHA/WHO FC</b>	n=2430	n=3642	n=850
I	203 (8.4)	59 (1.6)	30 (3.5)
II	1003 (41.3)	628 (17.2)	329 (38.7)
III	1116 (45.9)	2526 (69.4)	461 (54.2)
IV	108 (4.4)	429 (11.8)	30 (3.5)
<b>Systolic blood pressure mmHg</b>	n=2521	NA	n=342
<110	861 (34.1)		67 (19.6)
≥110	1660 (65.9)		275 (80.4)
<b>Heart rate bpm</b>	n=2523	NA	n=794
≤96	2138 (84.8)		680 (85.6)
>96	385 (15.2)		114 (14.4)
<b>6MWD m</b>	n=2212	n=2833	n=915
0–165	148 (5.9)	382 (13.5)	62 (6.8)
165–320	859 (34.0)	948 (33.5)	215 (23.5)
320–440	851 (33.7)	925 (32.7)	314 (34.3)
>440	671 (26.5)	578 (20.4)	324 (35.4)
<b>BNP/NT-proBNP pg·mL<sup>-1</sup></b>	n=1732	n=2883	NA
<50/<300	531 (30.7)	578 (20.0)	
50–200/300–1100	566 (32.7)	756 (26.2)	
200–800/NA	426 (24.6)	235 (8.2)	
≥800/≥1100	209 (12.1)	1314 (45.6)	
<b>D<sub>LCO</sub> % pred</b>	n=1625	n=1197	n=644
<40	373 (22.9)	424 (35.4)	218 (33.9)
≥40	1252 (77.1)	773 (64.6)	426 (66.2)
<b>RAP mmHg</b>	n=728	n=3141	n=955
≤20	709 (97.4)	3027 (96.4)	933 (97.7)
>20	19 (2.6)	114 (3.6)	22 (2.3)
<b>PVR WU</b>	n=2411	n=3094	n=874
<5	504 (20.9)	601 (19.4)	321 (36.7)
≥5	1907 (79.1)	2493 (80.6)	553 (63.3)
<b>eGFR mL·min<sup>-1</sup>·1.73 m<sup>-2</sup></b>	n=1077	n=3040	NA
<60	357 (33.2)	1150 (37.8)	
≥60	720 (66.8)	1890 (62.2)	
<b>Pericardial effusion</b>	n=2216	NA	n=970
Yes	565 (25.5)		135 (13.9)
No	1651 (74.5)		835 (86.1)
<b>Hospitalisation within 6 months</b>		NA	n=403
Yes	407 (16.1)		59 (14.6)
No	2122 (83.9)		344 (85.4)

Data are presented as n (%), unless otherwise stated. WHO: World Health Organization; CTD-PAH: connective tissue disease-associated pulmonary arterial hypertension; POPH-PAH: portopulmonary hypertension-associated pulmonary arterial hypertension; FPAH: familial pulmonary arterial hypertension; NYHA: New York Heart Association; FC: functional class; 6MWD: 6-min walk distance; BNP: brain natriuretic peptide; NT-proBNP: N-terminal pro-BNP; NA: not applicable; D<sub>LCO</sub>: diffusing capacity of the lung for carbon monoxide; RAP: right atrial pressure; PVR: pulmonary vascular resistance; WU: Wood Units; eGFR: estimated glomerular filtration rate.

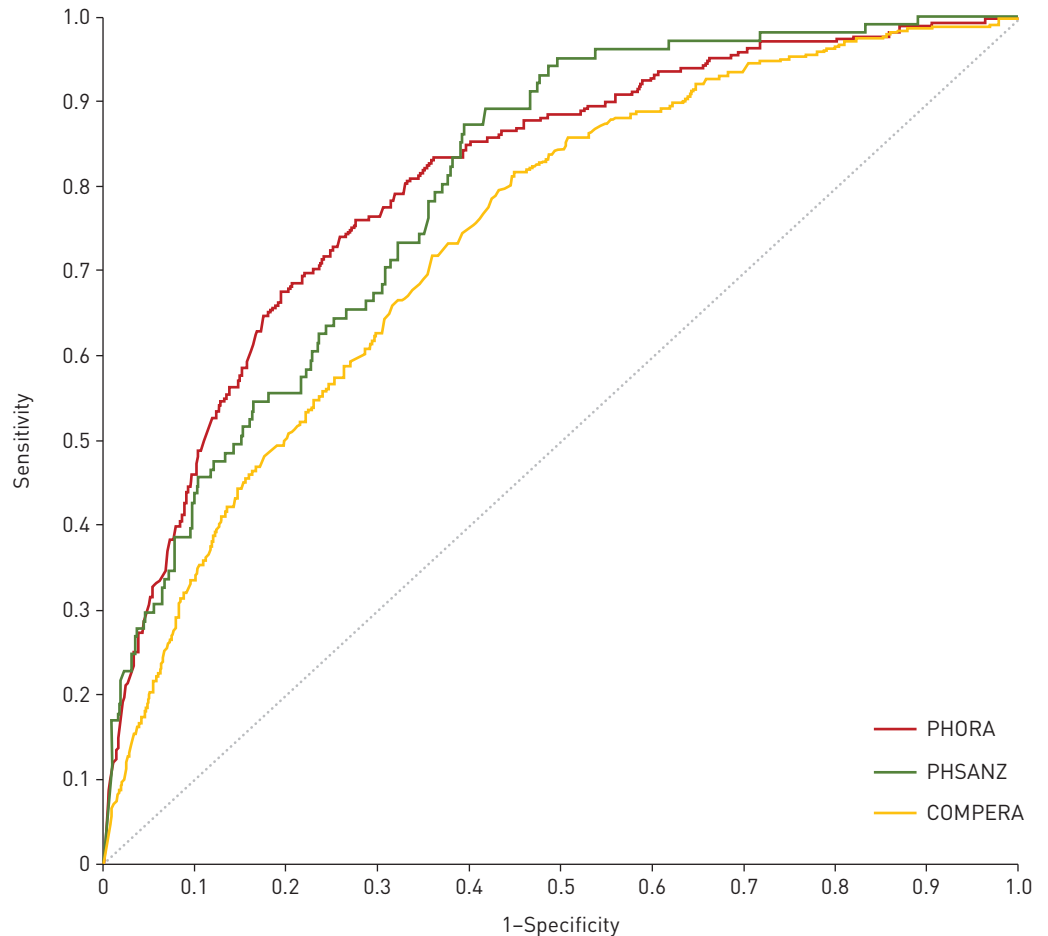


FIGURE 2 Performance of the Bayesian networks algorithm when internally validated in the Registry to Evaluate Early and Long-Term PAH Disease Management (REVEAL) (Pulmonary Hypertension Outcomes Risk Assessment (PHORA); area under the curve [AUC] 0.80), and externally in the Pulmonary Hypertension Society of Australia and New Zealand (PHSANZ; AUC 0.80) and Comparative Prospective Registry of Newly Initiated Therapies for Pulmonary Hypertension (COMPERA; AUC 0.74) registries.

survival rates predicted by PHORA were greater for patients with lower risk scores and poorer for those with higher risk scores ( $p < 0.001$ ), with excellent separation between low-, intermediate- and high-risk groups in all three registries (figure 3). This demonstrates PHORA's ability to risk-stratify patients effectively early in the course of the disease, which would allow for appropriate clinical decision making.

Figure 4 demonstrates the ability of PHORA to illustrate the dynamic interdependencies among the variables. Figure 4a demonstrates the baseline probability relationships between variables in the model and the outcome during a baseline assessment of an example patient. Figure 4b shows how these baseline probability relationships of the network change with the addition of new variables as patient undergoes ongoing work-up.

### Discussion

Risk stratification using a Bayesian network model approach (PHORA) provides improved discrimination to the existing Cox regression multivariate model (REVEAL 2.0), and effectively depicted risk in two large external registry cohorts, COMPERA and PHSANZ. This improvement probably stems from the ability of the Bayesian network model to understand both the dynamic influences of each risk factor on each other, as well as with the outcome itself.

The utility of the Bayesian network methodology was only recognised within the past 25 years, with the publication and application of Bayesian network-based decision support tools in a variety of medical disciplines [13–16]. In these clinical scenarios, Bayesian network-based tools were noted to have superior predictive performance over traditional statistical methods [8]. Bayesian networks do not require restrictive

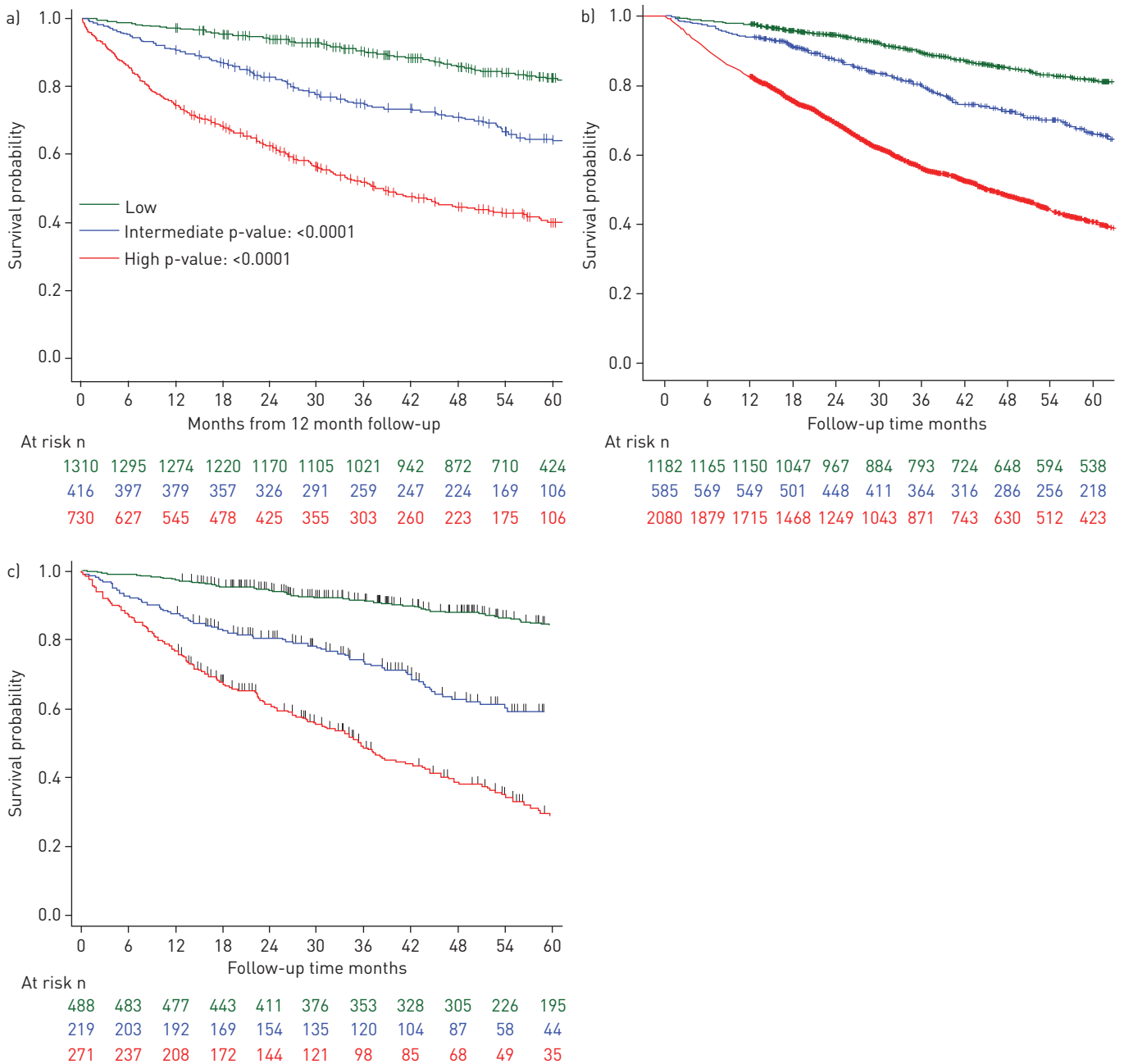


FIGURE 3 Kaplan-Meier curves demonstrating Pulmonary Hypertension Outcomes Risk Assessment (PHORA)'s risk-stratification abilities into low, intermediate and high risk of 12-month mortality based on the 2015 European Society of Cardiology/European Respiratory Society guidelines in a) the Registry to Evaluate Early and Long-Term PAH Disease Management (REVEAL); b) the Comparative Prospective Registry of Newly Initiated Therapies for Pulmonary Hypertension (COMPERA); and c) Pulmonary Hypertension Society of Australia and New Zealand (PHSANZ) registry.

modelling assumptions outside of expressing independencies whenever these are justified. Descriptively, Bayesian networks provide the advantages of a rigorous probabilistic framework that uses inference of multiple variables and a visual representation that is interactive and easy to interpret. This also allows a user to input these various scenarios and calculate the changes in predicted mortality and other adverse events in a highly interactive fashion. When performing prediction, Bayesian networks allow for estimating the outcome probability based on partial observations, as often happens in a clinical setting. Indeed, just converting the methodology of evaluating the pertinent REVEAL 2.0 variables produced a tool with boosted the discriminatory power of the model (from an AUC of 0.76 to an AUC of 0.80) [17]. Whether this improvement translates to clinical significance remains to be seen. Lastly, Bayesian networks offer more flexibility and result in more intuitive models.

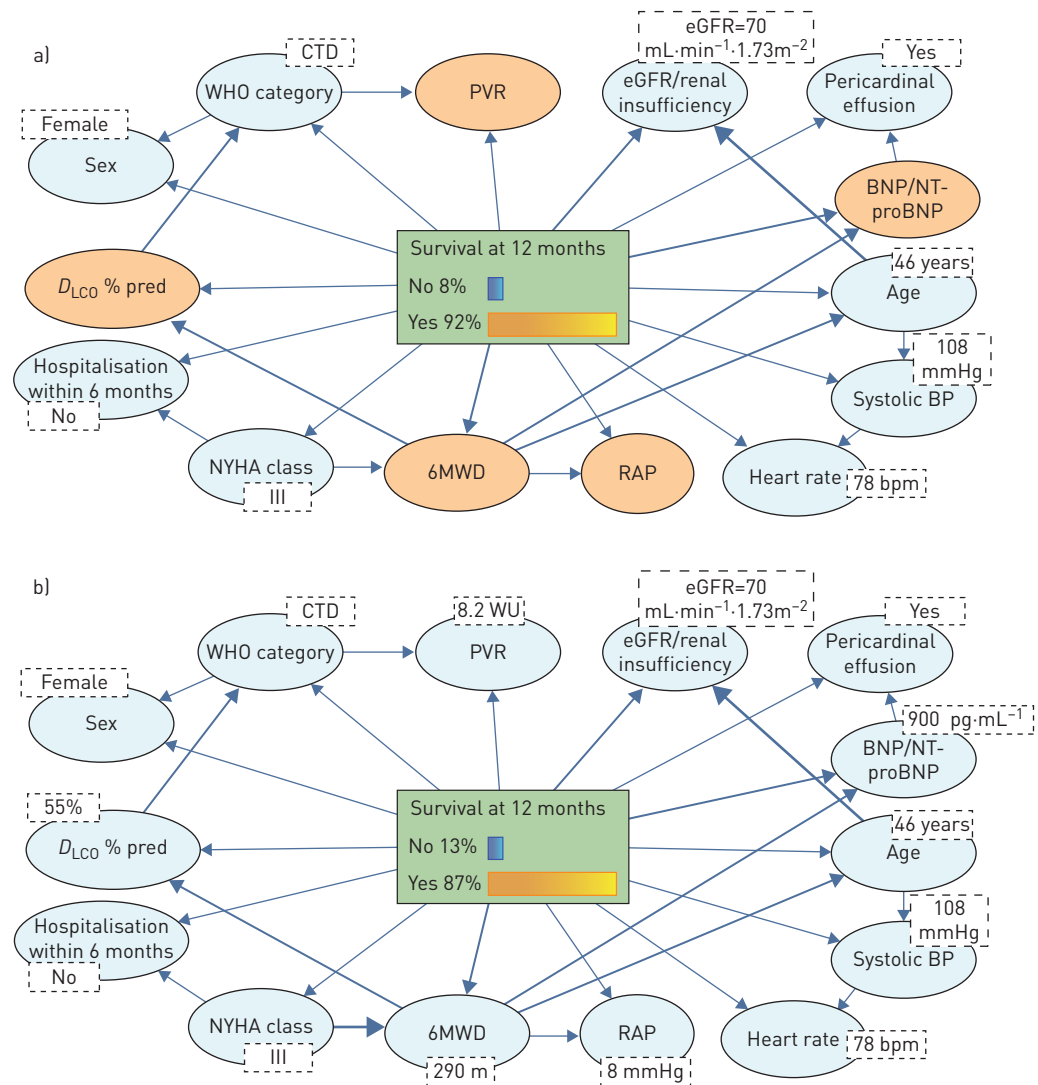


FIGURE 4 a) Example of a Pulmonary Hypertension Outcomes Risk Assessment (PHORA) model when some variables (highlighted in blue) are observed at baseline assessment. The values of these variables are noted in the dotted line box adjacent to each node. Variables in orange are yet to be reported as patient is undergoing work-up. b) Updated PHORA model when additional parameters (previously in orange) are now available. Note change in the predicted outcome [survival at 12 months, green box] as additional data is input. PVR: pulmonary vascular resistance; eGFR: estimated glomerular filtration rate; NT-proBNP: N-terminal pro-BNP; BP: blood pressure; RAP: right atrial pressure; 6MWD: 6-min walk distance; NYHA: New York Heart Association;  $D_{LCO}$ : diffusing capacity of the lung for carbon monoxide; WHO: World Health Organization; CTD: connective tissue disease.

Appropriate risk-stratification tools are necessary to guide clinical treatment goals and monitor disease progression. Clinically, a good risk assessment tool should be evidence based, easy to administer, externally validated, have good discrimination (C-index >0.7), account for “missingness” in data, incorporate weighting of individual variables and reflect the dynamic interactions between variables as well the primary outcome [2]. In the development of contemporary risk stratification in PAH, investigators are limited in their ability to produce robust and highly discriminatory (*i.e.* C-index >0.8) predictive tools. This relates in part reliance on registry datasets, which are limited in data quality, quantity and comprehensiveness. Although real-world in nature, these registries provide limited yield of high-quality data in light of the differences in patient characteristics enrolled, number of patients observed, quality of data collected and failure to capture relevant variables (*i.e.* imaging or novel biomarkers) that could add substantially to the comprehensiveness and discriminatory power of equations and calculator. Another significant limitation to the predictive power of contemporary risk assessments is their reliance on traditional statistical methods (Cox proportional hazard) or expert opinion. Cox proportional hazard models allow for estimating the effect of multiple risk factors on survival, with the impact of each



individual risk factor expressed by their hazard ratio. However, hazard ratio remains constant over time and is unaffected by concomitant risk factors [18]. In addition, clinically relevant variables such as rate of disease progression remain unaccounted for [19]. Lastly, traditional models are not capable of handling several missing clinical variables, which may not have been obtained at the time of evaluation. This results in a unidimensional and sometimes oversimplified risk prediction, which lacks in robustness with respect to predicting outcome in complex disease. Thus, at this point, until new datasets are made available, adapting our statistical methodology may improve our discrimination. The use of Bayesian networks could help with several of these shortcomings.

As per the 2015 ESC/ERS treatment guidelines, PAH should be risk-stratified as low (<5%), intermediate (5–10%) or high (>10%) risk of mortality at 12 months, to enable guidance on therapeutic decisions. However, in clinical practice, some patients may present with a combination of low-, intermediate- or high-risk features, which can then cloud clinical judgment and misguide subsequent medical therapy. PHORA can be deployed as a decision tool in the clinical arena to integrate the sometimes conflicting information. Another unique advantage of PHORA is that it allows for estimation of the outcome probability based on partial observations, without knowledge of presence or absence of remaining risk factors (figure 4).

Although PHORA was derived from a primarily prevalent patient registry (REVEAL), it was able to predict outcomes with equally good discrimination across two completely different real-world registries, regardless of whether patients were mostly incident (COMPERA) or prevalent (PHSANZ). Lastly, longitudinal monitoring with PHORA could guide treatment strategies by providing a specific, quantitative metric for satisfactory clinical response (a relative reduction of baseline percentage risk as opposed to lowering a risk stratum). It is envisioned that PHORA outputs and clinical variable entry will be depicted in an easy-to-visualise format on a web-based application, along with comparative REVEAL 2.0, COMPERA and French scores [5, 7] (www.myphora.org; figure 5), allowing a side-by-side decision tool

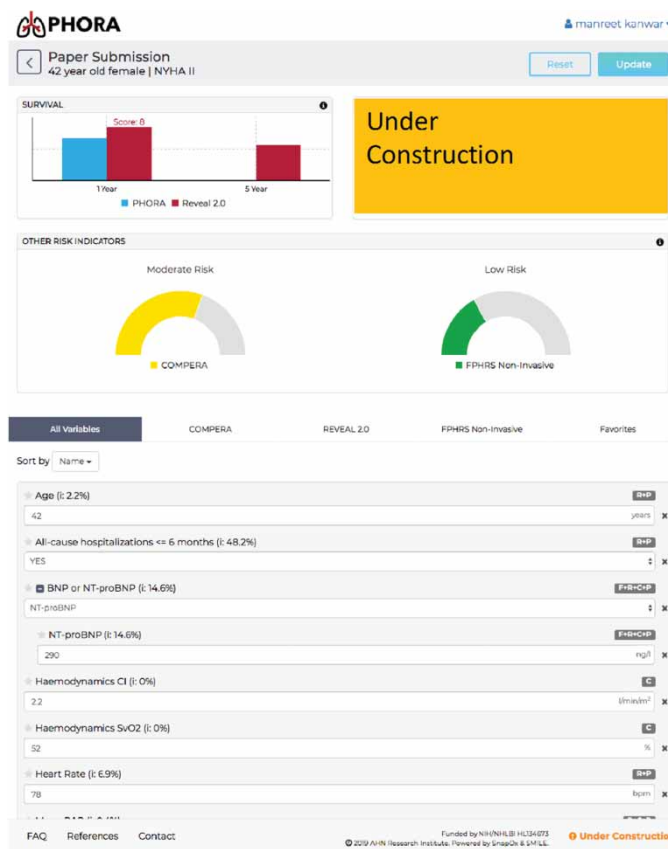


FIGURE 5 A screenshot of the webpage that will demonstrate the predicted clinical outcome (survival at 12 months). Outcomes as predicted by Pulmonary Hypertension Outcomes Risk Assessment (PHORA) are shown as a blue bar, as predicted by Registry to Evaluate Early and Long-Term PAH Disease Management (REVEAL) 2.0 as a red bar at 1 and 5 years; Comparative Prospective Registry of Newly Initiated Therapies for Pulmonary Hypertension (COMPERA) risk stratification is shown in yellow; and French noninvasive score in green. The clinical variables are shown at the bottom.

for clinicians to understand both the ranges in risk, the degree of influence of each variable on predicted outcome and likelihood scenario of each clinical case added.

We acknowledge that this study has several important limitations of deriving this new tool from clinical registry data, including missing data pertaining to the independent variables. Although the REVEAL database is large and representative, like other registries it suffers from incomplete capture of many data elements. This could impact the analysis by allowing patients used in both the model training and validation whom have up to 40% of their data missing. This could be particularly pertinent, if the missing data are related to the health of the patient per se (e.g. patient was too sick, so tests could not be done), thus skewing the analysis toward healthy patients. However, the fact that the model is not built on 'ideal/complete' datasets and can handle data missingness is also reflective of real-life clinical scenarios where all clinical data may not be available at each time-point. An additional limitation is the dependency on REVEAL-based cut-points and data used to derive PHORA only reflected prevalent patients who were alive and in the study at 12 months of follow-up. This was done to account for all-cause hospitalisation data in the previous 6 months, but raises concerns that the risk score is subject to survival bias. However, risk prognostication is typically not subject to survivor bias because risk is assessed only during the time the patient has participated in the registry. Whether a change in projected risk prediction scores in PAH reflects a true change in a patient's outcome remains a topic of debate. Lastly, interactions noted between the variables and survival are clinically likely to be even more complex than was captured by the TAN model.

In order to address these limitations, further derivation and validation studies using Bayesian networks that can appropriately handle mixed (categorical and continuous) data are already in progress in a harmonised, contemporary clinical trial dataset (n>3000) in conjunction with the United States Food and Drug Administration (FDA). A combination of both feature engineering (evidence-based, expert guided selection), feature learning (*via* information scoring) and dimensionality reduction (*via* unsupervised methods) will be incorporated in these newer iterations of PHORA with a key goal of maximising its discrimination (C-index >0.8), while keeping the tool easy to use. Newer versions of PHORA will not rely only on existing REVEAL variables, but will include other novel and significant variables determined by unsupervised modelling methods and further enhanced by expert opinion. Lastly, Bayesian network-based models at follow-up time-points will be evaluated to capture the impact of variables that may change over time allowing a more comprehensive prediction based on disease progression. We believe that such analyses will allow for a cumulative risk analysis, balancing therapy side-effects against improved outcomes in PAH patients. Moreover, we hope to be able to demonstrate a change in score in response to therapy as being reflective of improved survival in this analysis.

The FDA advocates the prospective use of patient characteristic(s) to select a study population in which detection of a drug effect (benefit, or lack thereof) is more likely than in an unselected population. The use of enhanced risk scores in PAH drug efficacy trials could accommodate enrolment of patients that are deemed to be at intermediate- or high-risk for clinical worsening, hence allowing for substantially smaller sample size and cost-saving.

### Conclusion

Our Bayesian network-derived risk prediction model, PHORA, demonstrated an improvement in discrimination over existing models. Bayesian network models have the advantage to learn from available data, incorporate expert knowledge, account for the interrelationships between clinical variables on outcome, and are more tolerant to missing data elements when calculating predictions. Hence machine learning based risk modelling can provide PAH clinicians with a greater level of confidence for making medical decisions in this complex, progressive disease.

Conflict of interest: M.K. Kanwar reports grants from NIH/NHLBI, during the conduct of the study. M. Gomberg-Maitland reports consultancy/steering committee, data monitoring board work for Acceleron, Actelion, Complexa, Gossamer Bio, Reata, and Neuroderm; George Washington School of Medicine and Health Sciences has received grants for research from Altavant and United Therapeutics; and is a member of the scientific advisory board for United Therapeutics, outside the submitted work. M. Hoepfer reports personal fees from Actelion, Bayer, MSD and Pfizer, outside the submitted work. C. Pausch has nothing to disclose. D. Pittrow reports personal fees from Actelion, Bayer, Amgen, Boehringer Ingelheim, Sanofi, MSD and Biogen, outside the submitted work. G. Strange reports grants from Actelion Pharmaceuticals, GlaxoSmithKline and Bayer Pharmaceuticals, during the conduct of the study. J.J. Anderson reports grants from GlaxoSmithKline, non-financial support from Actelion and Bayer, personal fees from AstraZeneca, outside the submitted work. C. Zhao is an employee of Actelion Pharmaceuticals US, Inc., a Janssen Pharmaceutical Company of Johnson & Johnson. J.V. Scott has nothing to disclose. M.J. Druzdzal is a partner at BayesFusion, LLC. J. Krajangka has nothing to disclose. L. Lohmueller has nothing to disclose. J. Antaki reports grants from NIH/NHLBI (R01 HL134673), during the conduct of the study. R.L. Benza reports grants from NIH/NHLBI (R01 HL134673), Actelion, United Therapeutics and Bayer, during the conduct of the study.

Support statement: Funding for this work was provided by National Institutes of Health, Division of National Heart, Lung, and Blood Institute grants R01 HL134673, PHORA: Pulmonary Hypertension Outcomes Risk Assessment. Funding information for this article has been deposited with the Crossref Funder Registry.

## References

- 1 Benza RL, Miller DP, Barst RJ, *et al.* An evaluation of long-term survival from time of diagnosis in pulmonary arterial hypertension from the REVEAL Registry. *Chest* 2012; 142: 448–456.
- 2 Benza RL, Farber HW, Selej M, *et al.* Assessing risk in pulmonary arterial hypertension: what we know, what we don't. *Eur Respir J* 2017; 50: 1701353.
- 3 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.
- 4 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.
- 5 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.
- 6 Boucly A, Weatherald J, Savale L, *et al.* Risk assessment, prognosis and guideline implementation in pulmonary arterial hypertension. *Eur Respir J* 2017; 50: 1700889.
- 7 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.
- 8 Kraisangka J, Druzdzal MJ, Lohmueller LC, *et al.* Bayesian network vs. Cox's proportional hazard model of PAH risk: a comparison. In: Riaño D, Wilk S, ten Teije A, eds. *Artificial Intelligence in Medicine*. Cham, Springer International Publishing, 2019.
- 9 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.
- 10 McGoon MD, Krichman A, Farber HW, *et al.* Design of the REVEAL registry for US patients with pulmonary arterial hypertension. *Mayo Clin Proc* 2008; 83: 923–931.
- 11 Friedman N, Geiger D, Goldszmidt M. Bayesian network classifiers. *Machine Learning* 1997; 29: 131–163.
- 12 Strange G, Lau EM, Giannoulatou E, *et al.* Survival of idiopathic pulmonary arterial hypertension patients in the modern era in Australia and New Zealand. *Heart Lung Circ* 2018; 27: 1368–1375.
- 13 Kanwar MK, Lohmueller LC, Kormos RL, *et al.* A Bayesian model to predict survival after left ventricular assist device implantation. *JACC Heart Fail* 2018; 6: 771–779.
- 14 Loghmanpour NA, Kormos RL, Kanwar MK, *et al.* A Bayesian model to predict right ventricular failure following left ventricular assist device therapy. *JACC Heart Fail* 2016; 4: 711–721.
- 15 Miranda E, Irwansyah E, Amelga AY, *et al.* Detection of cardiovascular disease risk's level for adults using naive Bayes classifier. *Healthc Inform Res* 2016; 22: 196–205.
- 16 Cao K, Xu J, Zhao WQ. Artificial intelligence on diabetic retinopathy diagnosis: an automatic classification method based on grey level co-occurrence matrix and naive Bayesian model. *Int J Ophthalmol* 2019; 12: 1158–1162.
- 17 Mandrekar JN. Receiver operating characteristic curve in diagnostic test assessment. *J Thorac Oncol* 2010; 5: 1315–1316.
- 18 Cox DR. Regression models and life-tables. *JR Stat Soc B* 1972; 34: 187–220.
- 19 Hernán MA. The hazards of hazard ratios. *Epidemiology* 2010; 21: 13–15.