Early View

Original article

Risk stratification in pulmonary arterial hypertension using Bayesian analysis

Manreet K Kanwar, Mardi Gomberg-Maitland, Marius Hoeper, Christine Pausch, David Pittow, Geoff Strange, James J Anderson, Carol Zhao, Jacqueline V Scott, Marek J Druzdzel, Jidapa Kraisangka, Lisa Lohmueller, James Antaki, Raymond L Benza

This manuscript has recently been accepted for publication in the *European Respiratory Journal*. It is published here in its accepted form prior to copyediting and typesetting by our production team. After these production processes are complete and the authors have approved the resulting proofs, the article will move to the latest issue of the ERJ online.

Title: Risk Stratification in Pulmonary Arterial Hypertension Using Bayesian Analysis

Authors: Manreet K Kanwar[1], Mardi Gomberg-Maitland[2], Marius Hoeper[3], Christine Pausch[4], David Pittow[5], Geoff Strange[6], James J Anderson[7], Carol Zhao[8], Jacqueline V Scott[9], Marek J Druzdzel[10], Jidapa Kraisangka[11,] Lisa Lohmueller[12], James Antaki[13]; Raymond L Benza[14]

Institutional Affiliations:

1 - Cardiovascular Institute at Allegheny Health Network, Pittsburgh, Pennsylvania, USA

2 - George Washington School of Medicine and Health Sciences, Washington DC, USA

3 - Department of Respiratory Medicine, Hannover Medical School, German Center for Lung Research (DZL), Hannover, Germany

4 - GWT-TUD GmbH, Dresden, Germany

5 – Faculty of Institute for Clinical Pharmacology, Technical University, Dresden, Germany

6 – School of Medicine, University of Notre Dame, Fremantle, Australia

7 – Respiratory Department, Sunshine Coast University Hospital, Queensland, Australia

8 – Actelion Pharmaceuticals US, Inc., A Janssen Pharmaceutical Company of Johnson & Johnson, San Francisco, California, USA

9 – School of Biomedical Engineering, Carnegie Mellon University, Pittsburgh, Pennsylvania, USA

10 - Faculty of Computer Science, Bialystok University of Technology, Bialystok, Poland

11 - Faculty of Information and Communication Technology, Mahidol University, Thailand

12 - Language Technologies Institute, Carnegie Mellon University, Pittsburgh, Pennsylvania, USA

13 – Meinig School of Biomedical Engineering, Cornell University, Ithaca, New York, USA

14 - Ohio State Medical Center, Columbus, Ohio, USA

Corresponding author:

Raymond L. Benza, MD

Address: 320 E North Avenue, Pittsburgh, Pennsylvania – 15212, USA

Tel: (412) 359-4760

Fax: (412) 359-6544

Email: reaymond.benza@osumc.org

"Take home" message:

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 2 independent PAH registries.

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 (BN) 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 one-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 COMPERA and PHSANZ registry). Patients were classified as low, intermediate and high-risk (<5%, 5-20% and > 10% 12-month mortality, respectively) based on the 2015 ESC/ERS guidelines.

Results: PHORA had an AUC of 0.80 for predicting one-year survival, which was an improvement over REVEAL 2.0 (AUC of 0.76). When validated in COMPERA and PHSANZ registries, PHORA demonstrated an AUC of 0.74 and 0.80 respectively. One-year survival rates predicted by PHORA were greater for patients with lower risk scores and poorer for those with higher risk scores (P<.001), with excellent separation between low-, intermediate-, and high-risk groups in all three registries.

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

Introduction

Pulmonary arterial hypertension (PAH) is a rapidly progressive, incurable disease with a median survival of approximately 7 years after diagnosis.(1) Accurate risk stratification in PAH accommodates demographic, clinical, hemodynamic, and functional parameters, allowing clinicians to identify treatment goals, monitor disease progression and facilitate timely referral to a PAH center 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 (FPHRS), 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 inter-variable relationships.

Bayesian networks (BN) are highly efficient and sophisticated algorithms derived using data mining - a process of discovering patterns in pre-existing data. A BN can be trained to recognize 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, non-linear 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 paper, we sought to demonstrate the utility of BN based machine learning in enhancing the predictive ability of an 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 73% 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 BN models were derived from the final study data of REVEAL 2.0 (9) and included PAH patients who survived ≥1-year post-enrolment to allow sufficient capture of all-cause hospitalization data in the previous 6 months (derivation cohort).

Model Development with Bayesian Networks:

BNs 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 one-year mark available were included, using variables at 12 month, if available. If there was no assessment done at one year, variable most recent to that time point (including assessment at enrollment, up to 12 months) were 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 discretized into pre-specified intervals (e.g. NT pro BNP levels [<300, 300-1100, >1100 pg/mL] or 6MWD [<165, 165-320, 320-440, >440 meters]) as is required for Bayesian methodology. The BN model learned the direction and magnitude of influence between these pre-specified variables on each other as well as the final clinical outcome, represented in the model as conditional probability tables (CPTs). 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 that provides a platform for artificial intelligence modeling based on Bayesian networks. (https://www.bayesfusion.com/) (23)

Patient Population/ Validation cohorts:

We validated the PHORA BN model both internally and externally, utilizing the following cohorts and methodologies:

<u>Internal Validation</u>: We validated the PHORA model internally within the REVEAL registry using 10-fold cross validation and report the results of this validation as AUC.

<u>External Validation</u>: We validated the PHORA model externally in two registries:

(i)   The COMPERA registry, which is an ongoing multi-national European registry comprised of PH/PAH patients enrolled since May 2007.(4) PHORA model was validated on 3,849 newly diagnosed, consecutively enrolled PAH patients. Data from time of enrollment was considered.

(ii)   The Pulmonary Hypertension Society of Australia and New Zealand (PHSANZ) Registry, which collects data from patients with all subgroups of PH since December 2011 from 16 Australian and two New Zealand centers.(12) PHORA was validated in those PAH patients who had one-year data available (978 of 1076). Variables included were at the time closest to one-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 PHORA's ability to stratify risk groups in each of the three registry populations was calculated using chi-square analysis (SPSS, IBM).

Results

Of the 3,515 patients enrolled in REVEAL, 2,529 were in the registry at 12 months after enrollment and included in the PHORA derivation model. Of these, 73.7% were

previously diagnosed (i.e., >3 months before enrollment) and 26.3% were newly diagnosed (i.e., ≤3 months before enrollment). The majority of the patients were female (80%), NYHA/WHO FC II (41.3%) or FC 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 BN (PHORA, Figure 1) improved the predictive power of the calculator. The AUC of 0.80 for predicting one-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 a 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 a positive predictive value of 0.97 [95% CI: 0.96 -0.98] for one-year survival. PHORA demonstrated an AUC of 0.74 and 0.80 when validated in the COMPERA and PHSANZ registry, 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. Twelve-month survival rates predicted by PHORA were greater for patients with lower risk scores and poorer for those with higher risk scores (*P*<.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 PHORA's ability to illustrate the dynamic interdependencies among the variables (Figure 4). 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 BN model approach (PHORA) provides improved discrimination to the existing cox regression multi-variate model, REVEAL 2.0. and effectively depicted risk in two large external registry cohorts, COMPERA and PHSANZ. This improvement likely stems from the BN model's ability to understand both the dynamic influences of each risk factor on each other, as well as with the outcome itself.

The utility of the BN methodology was only recognized within the past 25 years, with the publication and application of BN-based decision support tools in a variety of medical disciplines.(13-16) In these clinical scenarios, BN based tools were noted to have superior predictive performance over traditional statistical methods.(8) BNs do not require restrictive modeling assumptions outside of expressing independencies whenever these are justified. Descriptively, BNs 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, BNs 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, BNs offer more flexibility and result in more intuitive models.

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 CPH) or expert opinion. CPH 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 (HR). However, HR remains constant over time and unaffected by concomitant risk factors.(18) Also, 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 over simplified 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 upon our discrimination. The use of BNs 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. In clinical practice, however, 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 advantages 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 (COMEPRA) 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 % risk as opposed to lowering a risk strata). It is envisioned that PHORA outputs and clinical variable entry will be depicted in an easy to visualize format on a web-based application, along with comparative REVEAL 2.0, COMPERA and French scores (5,7) (www.myphora.org, see Figure 5), allowing a side by side decision tool 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 is 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 missing-ness 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 hospitalization 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 BNs that can appropriately handle mixed (categorical and continuous) data are already in progress in a harmonized, contemporary clinical trial dataset (N > 3000) in conjunction with the United States Food and Drug Association (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 maximizing 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 modeling methods and further enhanced by expert opinion. Lastly, BN-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 enrollment 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 BN derived risk prediction model, PHORA, demonstrated an improvement in discrimination over existing models. BN 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 modeling can provide PAH clinicians with a greater level of confidence for making medical decisions in this complex, progressive disease.

Conflict of interest

References:

1.   Benza RL, Miller DP, Barst RJ, Badesch DB, Frost AE, McGoon MD. An evaluation of long-term survival from time of diagnosis in pulmonary arterial hypertension from the REVEAL Registry. Chest. 2012;142(2):448-56.
2.   Benza RL, Farber HW, Selej M, Gomberg-Maitland M. Assessing risk in pulmonary arterial hypertension: what we know, what we don't. Eur Respir J. 2017;50(2).
3.   Galie N, Humbert M, Vachiery JL, Gibbs S, Lang I, Torbicki A, 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(4):903-75.
4.   Hoeper MM, Kramer T, Pan Z, Eichstaedt CA, Spiesshoefer J, Benjamin N, et al. Mortality in pulmonary arterial hypertension: prediction by the 2015 European pulmonary hypertension guidelines risk stratification model. Eur Respir J. 2017;50(2).
5.   D'Alonzo GE, Barst RJ, Ayres SM, Bergofsky EH, Brundage BH, Detre KM, et al. Survival in patients with primary pulmonary hypertension. Results from a national prospective registry. Annals of internal medicine. 1991;115(5):343-9.
6.   Boucly A, Weatherald J, Savale L, Jais X, Cottin V, Prevot G, et al. Risk assessment, prognosis and guideline implementation in pulmonary arterial hypertension. Eur Respir J. 2017;50(2).
7.   Kylhammar D, Kjellstrom B, Hjalmarsson C, Jansson K, Nisell M, Soderberg S, et al. A comprehensive risk stratification at early follow-up determines prognosis in pulmonary arterial hypertension. Eur Heart J. 2018;39(47):4175-81.

8. Kraisangka J, Druzdzel MJ, Lohmueller LC, Kanwar MK, Antaki JF, Benza RL, editors. Bayesian Network vs. Cox's Proportional Hazard Model of PAH Risk: A Comparison2019; Cham: Springer International Publishing.

9. Benza RL, Gomberg-Maitland M, Elliott CG, Farber HW, Foreman AJ, Frost AE, 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(2):323-37.

10. McGoon MD, Krichman A, Farber HW, Barst RJ, Raskob GE, Liou TG, et al. Design of the REVEAL registry for US patients with pulmonary arterial hypertension. Mayo Clin Proc. 2008;83(8):923-31.

11. Friedman N, Geiger D, Goldszmidt M. Bayesian Network Classifiers. Machine Learning. 1997;29(2):131-63.

12. Strange G, Lau EM, Giannoulatou E, Corrigan C, Kotlyar E, Kermeen F, et al. Survival of Idiopathic Pulmonary Arterial Hypertension Patients in the Modern Era in Australia and New Zealand. Heart Lung Circ. 2018;27(11):1368-75.

13. Kanwar MK, Lohmueller LC, Kormos RL, Teuteberg JJ, Rogers JG, Lindenfeld J, et al. A Bayesian Model to Predict Survival After Left Ventricular Assist Device Implantation. JACC Heart failure. 2018.

14. Loghmanpour NA, Kormos RL, Kanwar MK, Teuteberg JJ, Murali S, Antaki JF. A Bayesian Model to Predict Right Ventricular Failure Following Left Ventricular Assist Device Therapy. JACC Heart failure. 2016;4(9):711-21.

15. Miranda E, Irwansyah E, Amelga AY, Maribondang MM, Salim M. Detection of Cardiovascular Disease Risk's Level for Adults Using Naive Bayes Classifier. Healthc Inform Res. 2016;22(3):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(7):1158-62.

17. Mandrekar JN. Receiver operating characteristic curve in diagnostic test assessment. J Thorac Oncol. 2010;5(9):1315-6.

18. Cox DR. Regression Models and Life-Tables. Journal of the Royal Statistical Society Series B (Methodological). 1972;34(2):187-220.

19. Hernán MA. The hazards of hazard ratios. Epidemiology (Cambridge, Mass). 2010;21(1):13-5.

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 calculator

WHO, World Health Organization; PAH, Pulmonary Arterial Hypertension; APAH-portal, Associated PAH with Portal Hypertension; APAH- CVD/CTD, Associated PAH with Collagen Vascular Disease/Connective Tissue Disease; FPAH, Familial PAH; Other, includes Idiopathic PAH; eGFR, Estimated Glomerular Filtration Rate; NYHA FC, New York Heart Association Functional Class; NT-proBNP, N-terminal pro b-type Natriuretic Peptide; RAP, Right atrial pressure; DLCO, Diffusing capacity of the lungs for carbon monoxide; PVR, Pulmonary vascular resistance

Table 2: Clinical variables through 3 pulmonary arterial hypertension registries: REVEAL, Comparative, Prospective Registry of Newly Initiated Therapies for Pulmonary Hypertension (COMPERA) and Pulmonary Hypertension Society of Australia and New Zealand (PHSANZ) registry.

Figure 1: Structure of the Pulmonary Hypertension Outcomes Risk Assessment (PHORA) Bayesian network (BN) model, with conditional probability table (CPT) for survival

Figure 2: Performance of the BN algorithm when internally validated in REVEAL (PHORA, AUC 0.80), and externally in PHSANZ (AUC 0.80) and COMPERA (AUC 0.74) registries

Figure 3: Kaplan-Meier curves demonstrating PHORA's risk stratification abilities into low, intermediate and high risk of 12-month mortality based on the 2015 ESC/ERS guidelines in REVEAL (A), COMPERA (B), and PHSANZ (C) registries

Figure 4 (A) Example of a 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.
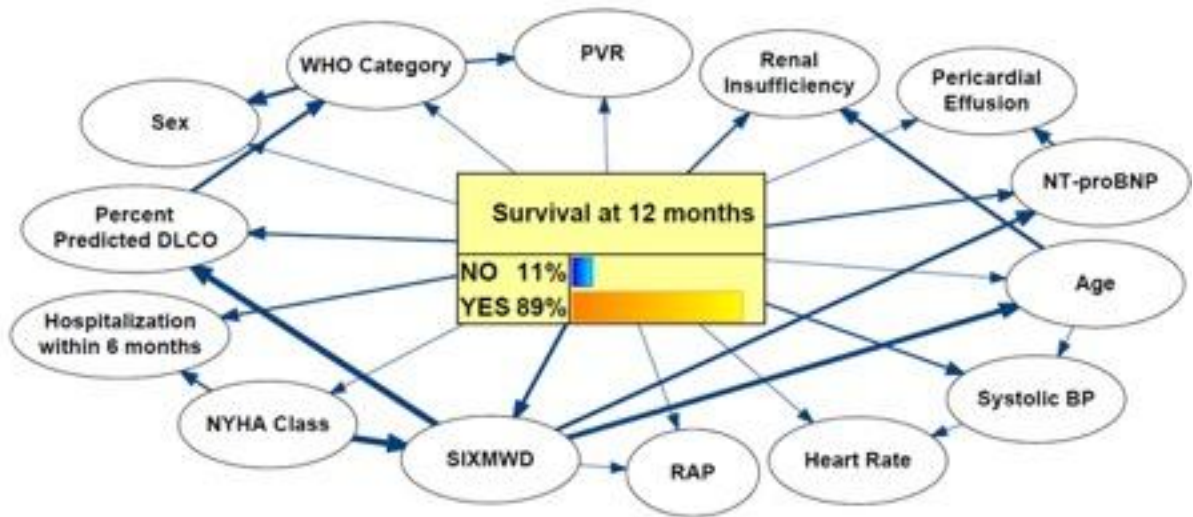
Figure 5 A screenshot of the webpage that will demonstrate the predicted clinical outcome (survival at 12 months). Outcomes as predicted by PHORA are shown as a blue bar, as predicted by REVEAL 2.0 as a red bar at 1 and 5 years, COMPERA risk stratification is shown in yellow and French non-invasive score as green. The clinical variables are shown at the bottom
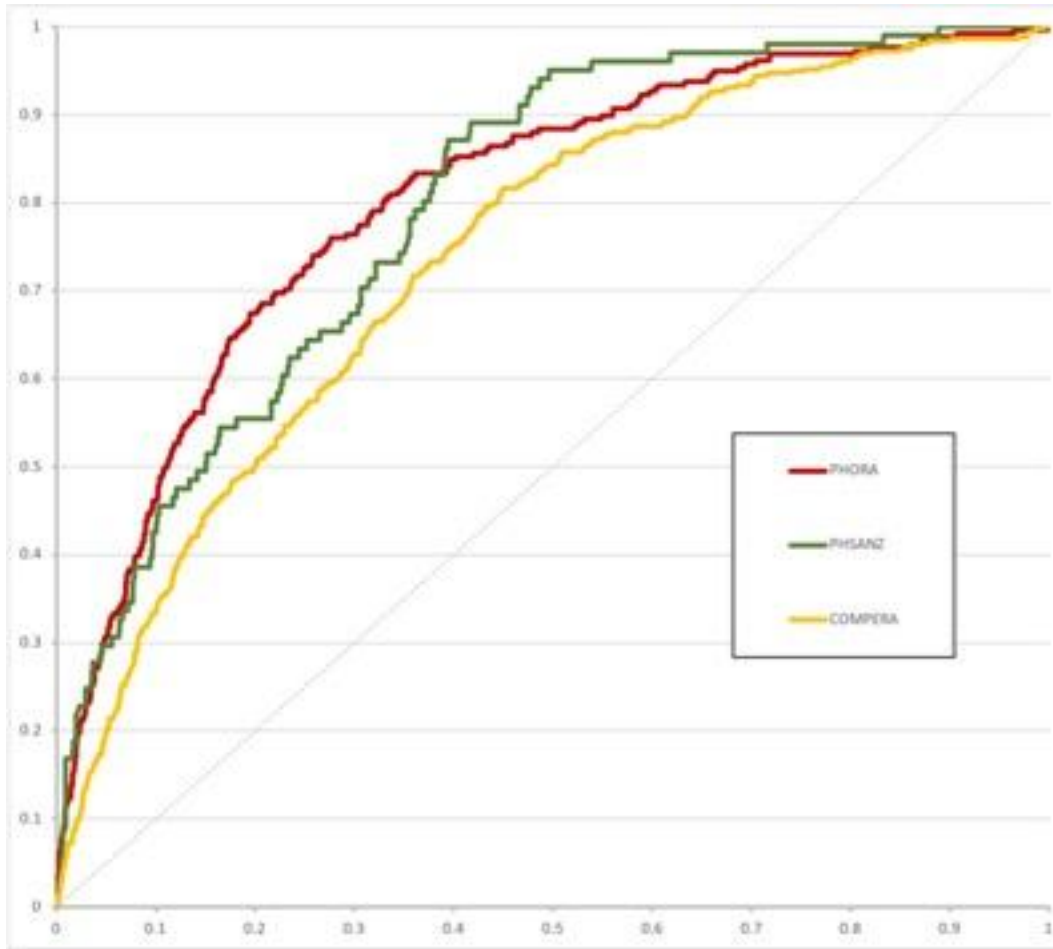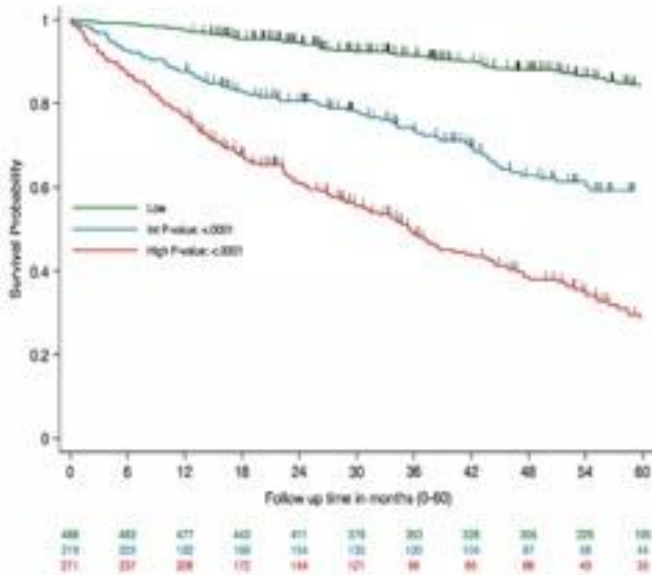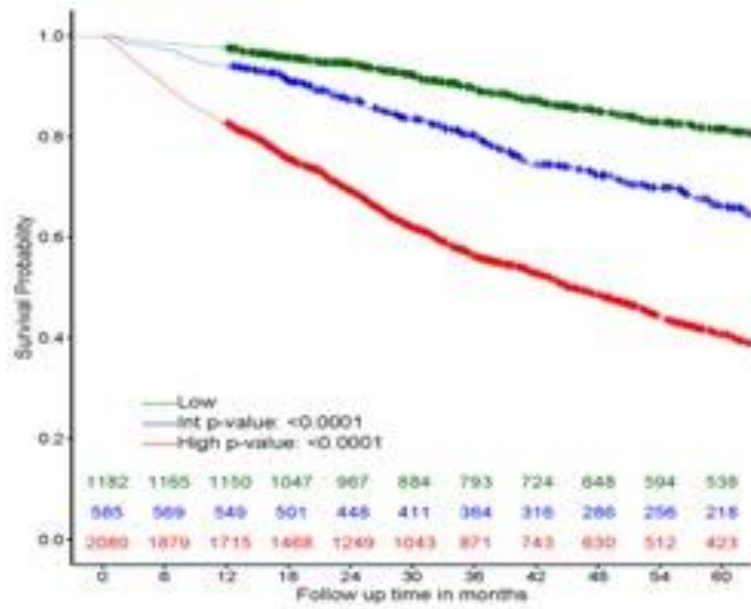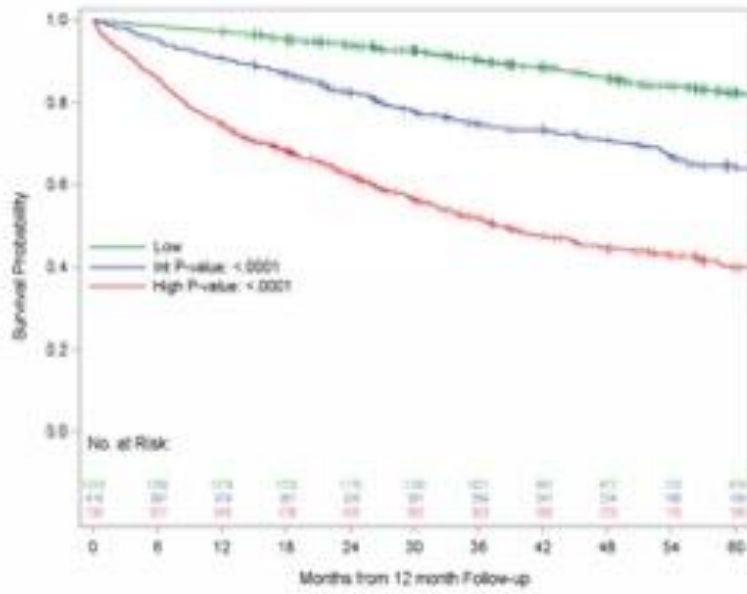
Table 1

Risk factor in REVEAL	Random Variable	Nodes in BN
CTD-PAH Heritable PoPH	WHO group I	CTD
		Heritable
		PoPH
		Other
Male > 60 years	Gender	Female
		Male
	Age (years)	≤60
		>60
Comorbidity eGFR < 60mL/min/1.73m2 or renal insufficiency if eGFR is unavailable	eGFR < 60 mL/min/1.73m2 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 BP < 110 mm Hg	Systolic BP (mm Hg)	<110
		≥110
Heart rate > 96 bpm	Heart Rate (bpm)	≤96
		>96
6 MWD < 165 meters 6MWD 320-440 meters 6MWD ≥ 440 meters	6MWD in meters	<165
		165 to <320
		320 to 440
		≥ 440
BNP < 50 pg/ml or NT-proBNP < 300 pg/mL BNP 200 to < 800 mg/mL BNP ≥ 800 pg/mL or NT pro BNP ≥ 1100 pg/mL	BNP or NT pro BNP (pg/mL)	<50 or <300
		50 – 200 or 300 -1100
		200 - 800
		≥800 or ≥1100
Echocardiogram Pericardial effusion	Pericardial effusion	Yes
		No
Right heart catheterization Mean RAP ≥ 20 mmHg within 1 year	Mean RAP	<20
PVR < 5 WU		≥20
	PVR in WU	< 5
		≥ 5
Pulmonary function test % DLCO <40%	% DLCO	<40
		≥40
Hospitalization in 6 months	Hospitalization in 6 months	Yes
		No
Survival at 12 months	Survival at 12 months	Yes
		No

Table 2

	REVEAL (n=2,529)	COMPERA (n=3,849)	PHSANZ (n=978)
Sex	n=2,529	n=3,849	n=978
• Male	505 (20.0%)	1,373 (35.7%)	218 (22.3%)
• Female	2,024 (80.0%)	2,476 (64.3%)	760 (77.7%)
Age (years)	n=2,529	n=3,849	n=978
• < 60	1673 (66.1%)	1,661 (43.2%)	392 (40.1%)
• > 60	856 (33.9%)	2,188 (56.8%)	586 (59.9%)
WHO Category	n=2,529	n=3,849	n=978
• Idiopathic	1,171 (46.3%)	2,158 (46.3%)	468 (47.9%)
• APAH-CTD	649 (25.7%)	1,347 (35.0%)	315 (32.2%)
• APAH-PoPH	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%)
NYHA/ WHO FC	n=2,430	n=3,642	n=850
• I	203 (8.4%)	59 (1.6%)	30 (3.5%)
• II	1,003 (41.3%)	628 (17.2%)	329 (38.7%)
• III	1,116 (45.9%)	2,526 (69.4%)	461 (54.2%)
• IV	108 (4.4%)	429 (11.8%)	30 (3.5%)
Systolic Blood Pressure (mm Hg)	n=2,521	N/A	n=342
• <110	861 (34.1%)		67 (19.6%)
• ≥110	1,660 (65.9%)		275 (80.4%)
Heart rate (per min)	n=2,523	N/A	n=794
• ≤96	2,138 (84.8%)		680 (85.6%)
• >96	385 (15.2%)		114 (14.4%)
6MWD (meters)	n=2,212	n=2,833	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%)
BNP/ NT-proBNP (pg/mL)	n=1,732	n=2,883	N/A
• <50/<300	531 (30.7%)	578 (20.0%)	
• 50-200/300-1100	566 (32.7%)	756 (26.2%)	
• 200-800/-	426 (24.6%)	235 (8.2%)	
• ≥800/≥1100	209 (12.1%)	1,314 (45.6%)	
DLCO (%)	n=1,625	n=1,197	n=644
• <40	373 (22.9%)	424 (35.4%)	218 (33.9%)
• ≥40	1,252 (77.1%)	773 (64.6%)	426 (66.2%)
RAP (mm Hg)	n=728	n=3,141	n=955
• ≤20	709 (97.4%)	3027 (96.4%)	933 (97.7%)
• >20	19 (2.6%)	114 (3.6%)	22 (2.3%)
PVR (Wood Units)	n=2,411	n=3,094	n=874
• <5	504 (20.9%)	601 (19.4%)	321 (36.7%)
• ≥5	1907 (79.1%)	2493 (80.6%)	553 (63.3%)
eGFR (mL/min/1.73m^2)	n=1,077	n=3,040	N/A
• <60	357 (33.2%)	1,150 (37.8%)	
• ≥60	720 (66.8%)	1,890 (62.2%)	
Pericardial Effusion	n=2,216	N/A	n=970
• Yes	565 (25.5%)		135 (13.9%)
• No	1,651 (74.5%)		835 (86.1%)
Hospitalization within 6 months	n=2,529	N/A	n=403
• Yes	407 (16.1%)		59 (14.6%)
• No	2,122 (83.9%)		344 (85.4%)

Paper Submission
42 year old female | NYHA II

Reset   Update

SURVIVAL

Score 8

1 Year   5 Year
■ PHORA ■ Reveal 2.0

Under
Construction

OTHER RISK INDICATORS

Moderate Risk

■ COMPERA

Low Risk
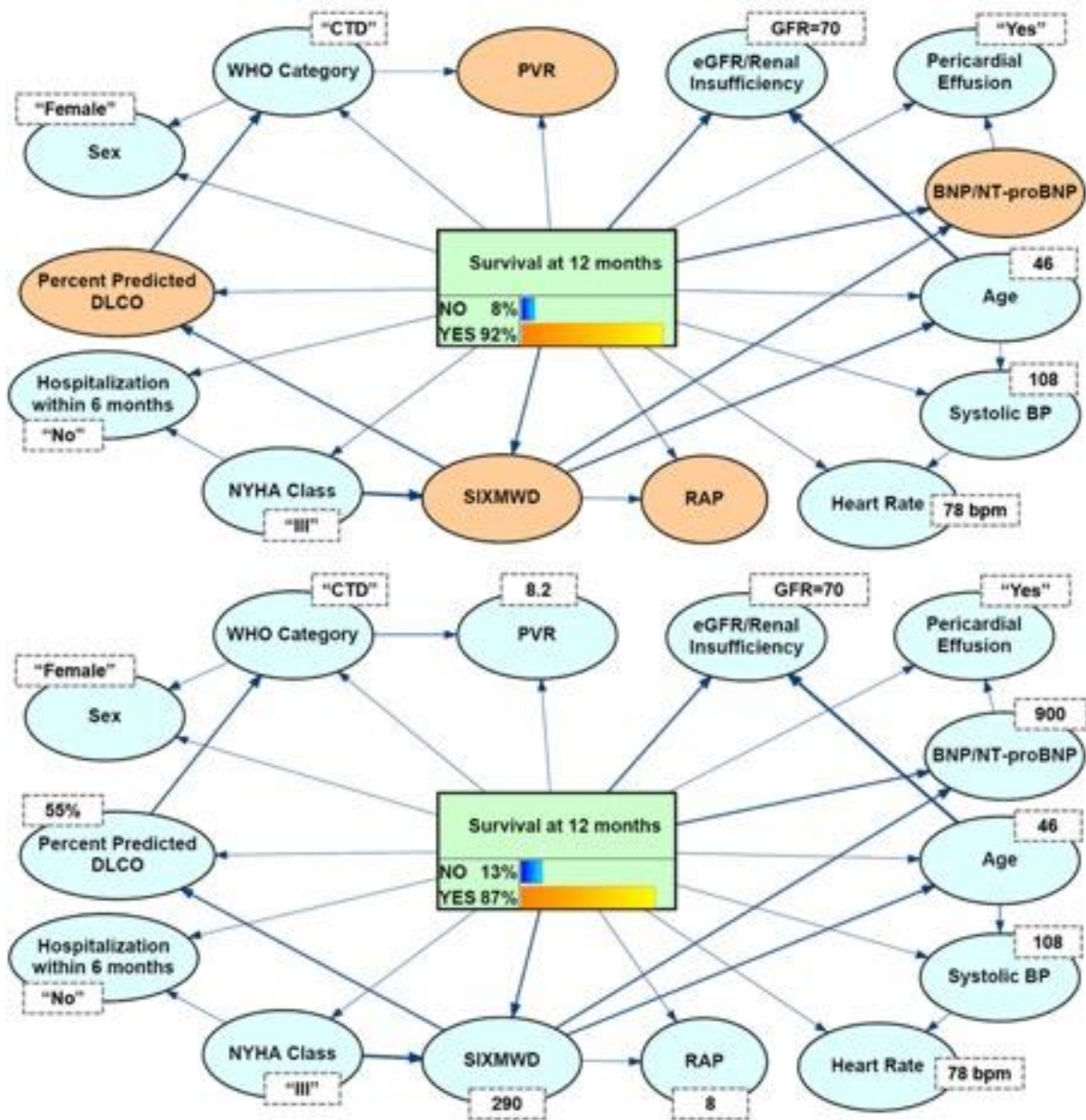
■ FPHRS Non-invasive

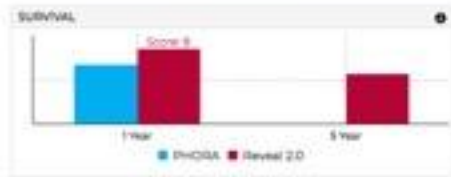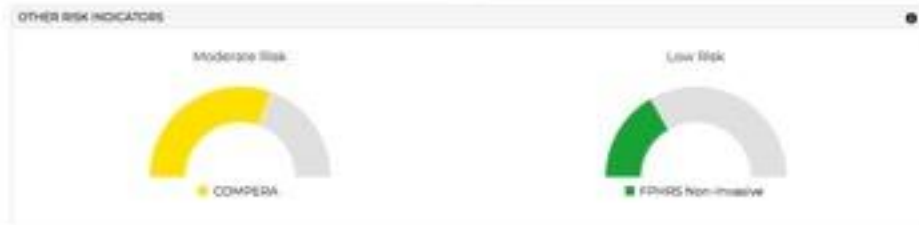| All Variables | COMPERA | REVEAL 2.0 | FPHRS Non-invasive | Favorites |

Sort by Name ▾

Age (: 2.2%)   BNP

42   years ✕

All-cause hospitalizations <= 6 months (: 48.2%)   BNP

YES   : ✕

■ BNP or NT-proBNP (: 14.6%)   BNP+CHF

NT-proBNP   : ✕

NT-proBNP (: 14.6%)   BNP+CHF

290   ng/l ✕

Haemodynamics CI (: 0%)   C

2.2   l/min/m² ✕

Haemodynamics SvO2 (: 0%)   C

52   % ✕

Heart Rate (: 6.9%)   BNP

78   bpm ✕

FAQ   References   Contact

Funded by NIH/NHLBI HL35875
© 2019 AHN Research Institute. Powered by SnapDx & DMCL.
● Under Construction