ABSTRACT FORM DEADLINE: December 15, 2021

Association of University Cardiologists
January 19-21, 2022
Belmond Charleston Place
Charleston, SC

Raymond Benza, MD, Ohio State University

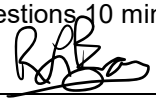
Background: Pulmonary arterial hypertension (PAH) is a fatal and difficult to treat disease due to patient inter-variability. Accurate patient risk stratification is necessary for guiding treatment, but no PAH risk calculator has been able to achieve an excellent performance in receiver-operator area under the curve (AUC > 0.8). New cutting-edge techniques with Bayesian network modeling for risk stratification have been published recently, show promising results (Pulmonary Hypertension Outcomes Risk Assessment or PHORA 1.0). Early studies of a newly derived Bayesian Network model (PHORA 2.0) demonstrated feasibility of improving PHORA 1.0 accuracy with an expanded feature set and new network structure. In this study, a finalized PHORA 2.0 model was developed that balances the number of variables required with model accuracy.

Methods: Patient-level data had been previously aggregated and harmonized across six contemporary PAH clinical trials (AMBITION, PATENT-1/2, GRIPHON, SERAPHIN, FREEDOM-EV, ARIES-1/2); all patients were assessed at baseline (for cross-over placebo patients from PATENT and ARIES, baseline was start of extension). Forty-one clinical variables were initially considered, based on their p-value ranking from previous meta-analyses, availability across trials, and expert opinion. A correlation heatmap was used to aggressively remove variables with moderate-to-strong correlation (R > 0.6), with priority given to the most significant variables in the meta-analysis. Training data was created by random sampling of 80% of the harmonized dataset, dropping early censored patients (N = 2531), leaving 20% of the data as a validation set (N = 626). Continuous variables were discretized through univariate supervised decision trees using 10-fold cross-validation that maximized Brier score. A novel method of genetic search optimization determined candidate groups of features that maximized ranked correlation with the outcome (Kendall's tau for one-year survival) and reduced redundant features, with increasing penalty for redundancy. This method established four feature combinations that were evaluated in Augmented Naïve Bayesian Network classifiers, and the best model was selected by multiple rounds of 10-fold cross-validation on training data. Final performance is reported as performance on the validation set.

Results: The best final model as determined from genetic search of feature combinations maintained a high cross-validation (Average AUC 0.82) using 16 out of 19 possible variables (see Figure 1). Final performance on the test set using this model was an AUC = 0.85. The final model outperformed all other published risk calculators at test time (Figure 1).

Conclusion: Bayesian network modeling coupled with genetic feature selection has discovered, for the first time, a mortality model for PAH with excellent performance (AUC > 0.8).

1) Please identify members by underlining their name.
2) Please use box above, Abstract (with spaces) = 500 Word limit
3) Talk duration 15 min, questions 10 min (total time 25 min)



Member's Signature

NOTE: This form is also available on the AUC website at https://www.aucard.org/scientific-abstract-form