SB^{3}C2024
Summer Biomechanics, Bioengineering and Biotransport Conference
June 11-14, Lake Geneva, WI, USA

A MACHINE LEARNING APPROACH TO MINING HEMODYNAMIC DATA FROM PULMONARY ARTERIAL HYPERTENSION RATS

Jingwen Hui (1), Ethan D. Kwan (1), Daniela Valdez-Jasso (1)

(1) Shu Chien – Gene Lay Department of Bioengineering, University of California San Diego, La Jolla, California, USA

INTRODUCTION

Pulmonary arterial hypertension (PAH) is characterized by sustained elevated blood pressure in the pulmonary arteries, causing a pressure overload on the right ventricle (RV). The classification and management of PAH is challenging because highly invasive procedures such as cardiac catheterization are required to confirm and track disease progression. Recent studies have described distinct patient remodeling phenotypes and therapy responses, revealing important differences in cardiac structure and function despite hemodynamic similarities [1],[2]. However, patient studies often face limitations in feasibility, invasiveness, and experimental design resulting in insufficient data relative to animal studies.

Our research team has previously obtained invasive measurements of cardiac structure and function involving large cohorts of PAH-induced male, female, and ovariectomized rats. Employing machine learning (ML), we leverage its predictive features and capabilities to explore underlying patterns challenging to discern through traditional methods. To that end, we 1) develop a predictive model using a support vector machine (SVM) with principal component analysis (PCA) for dimensionality reduction, 2) apply SVM and permutation importance analysis to understand quantitative differences across categories, and 3) investigate unexplored parameter relationships relevant to PAH through bivariate data visualization.

METHODS

Data acquisition. Hemodynamic and morphological data were collected from sugen-hypoxia (SuHx) treated rats as previously described [3]. Right-ventricular (RV) and left-ventricular (LV) blood pressure and volume were collected during an open chest procedure using a 1.9F admittance catheter (Transonic, Ithaca, NY) during steady state and vena cava occlusion. All additional hemodynamics metrics (stroke volume, end-diastolic volume, end-systolic volume, ejection fraction, cardiac output, dp/dt max, dp/dt min, pulmonary vascular resistance, end-diastolic elastance, end-systolic elastance, arterial elastance, ventricular vascular coupling) were derived from these timeseries. Morphological measurements (RV thickness, RV mass, Fulton index, animal mass, right/left atrial mass, septal hypertrophy index, liver mass) were collected immediately post hemodynamic studies. Septal Hypertrophy Index is calculated as LV mass divided by the sum of LV mass and septum mass, expressing the relative proportions of the LV chamber wall that make up the septum.

Numerical methods. SuHx animals were categorized as hypertensive if the mean pulmonary arterial pressure (mPAP) exceeds 20mmHg. If mPAP is unavailable, we defined hypertensive subjects as those that exceed 30mmHg for RV end-systolic pressure. Supervised ML, unsupervised ML, permutation importance analysis, and exploratory data visualizations were performed on the dataset comprising nearly equal proportions of male, female, and ovariectomized rats after 4, 8, 12, and 15 weeks of SuHx. Control group is referred to as week 0.

Weeks of SuHx/PAH Predictor

To investigate defining variables from this dataset, we performed PCA and fitted our dataset with a SVM model of Radial Basis Function (RBF) kernel. This model was chosen due to its effectiveness for handling high dimensional data with non-linear boundaries. Data was split into training and testing datasets (80%, 20%). Due to data imputation challenges, we only include RV hemodynamics and morphology data for this task.

Quantitative Categorical Analyzer

Of the rats treated with SuHx, some did not develop PAH. We define this group as non-hypertensive SuHx rats. First, we compared non-hypertensive SuHx *vs.* control rats using SVM and permutation importance analysis. Then we compared non-hypertensive SuHx to

hypertensive SuHx rats (i.e., treated *vs.* untreated SuHx). Through permutation importance analysis, we identify quantitative distinctions among these groups. Permutation importance coefficients quantify the change in our model's performance when the value of a variable is randomly shuffled, indicating the variable's contribution to the model's predictive power.

PAH Parameter Exploration

Leveraging pair plots, we explore correlations between morphological and hemodynamics parameters.

RESULTS

Weeks of SuHx/PAH Predictor

The model predicted PAH stage with 70% accuracy based on the weeks of SuHx. PC1, the most significant component, explains ~30% of the variance, which is more than double to that of PC2. The three variables with the greatest positive loadings of PC1 are RV mass, pulmonary vascular resistance (PVR), and end-systolic pressure (ESP). Conversely, the variables with the highest negative loadings are dp/dt min, ejection fraction, and cardiac output.

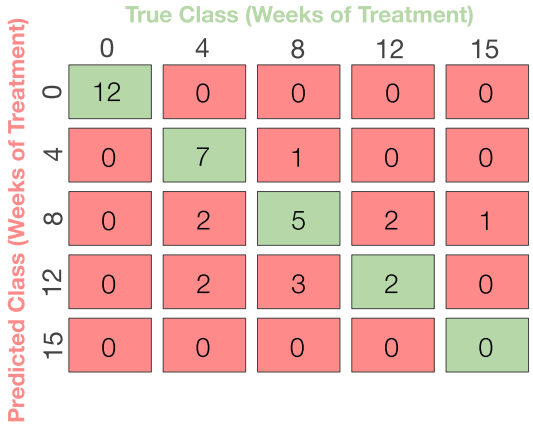

Figure 1: Confusion matrix of SVM shows significant accuracy in early-stage rats (week 4).

Quantitative Categorical Analyzer

The SVM model identifies control rats from non-hypertensive SuHx rats with a 92% accuracy level. The feature permutation analysis shows that the right atrial mass, left atrial mass, and ESP are the most significant indicators with permutation importance of 0.133, 0.075, and 0.036, respectively. On the other hand, SVM distinguishes hypertensive SuHx rats from the non-hypertensive SuHx ones with 100% accuracy. Feature permutation importance identified Fulton index, RV thickness, and RV mass as the key variables with permutation importance of 0.066, 0.039, 0.027, respectively.

$$\text{Permutation Importance} = \text{Original Accuracy} - \text{Permutated Accuracy} \quad (1)$$

$$\text{Accuracy} = \#\text{ of Correct Predictions} / \text{Total } \#\text{ of Predictions} \quad (2)$$

PAH Parameter Exploration

Pair plot visualizations showed previously unidentified relationships between Fulton index and three hemodynamics variables (dp/dt max, PVR, and end-systolic elastance (Ees)). When the Fulton index is low, these metrics tend to remain low. However, when the Fulton index is high, these parameters can vary across a range from low to high. The Pearson coefficients for the Fulton Index with dp/dt max, PVR, and Ees are 0.5878, 0.6033, and 0.4321, respectively.

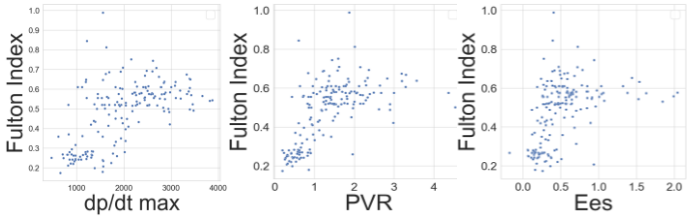

Figure 2: Strong linear correlations between hemodynamics parameters dp/dt max (left), PVR (middle), and Ees (right) and Fulton Index.

DISCUSSION

This study employed ML and data visualization tools to identify defining variables specific to different groups and uncover inter-categorical correlations. The SVM demonstrated exceptional accuracy in the early weeks, needing further fine-tuning to enhance prediction at the later weeks (weeks 8 and 12). Despite the limitations of a dataset with diverse y variables (weeks of treatment), the 70% accuracy achieved is noteworthy, particularly given previous SVM work mainly addressed binary y variables (hypertensive *vs.* normotensive).

The model also revealed distinctive characteristics of non-hypertensive SuHx rats. The right and left atrial masses were lower in non-hypertensive SuHx rats relative to hypertensive SuHx rats. Also, hypertensive SuHx animals displayed elevated values for the Fulton index and RV thickness compared to non-hypertensive SuHx rats, confirming increased right ventricular remodeling as hypothesized.

Also, the metric exploration study revealed correlations between non-invasive measurement of the Fulton Index with key hemodynamics metrics such as dp/dt max, PVR, and Ees. These three parameters are clinically collected invasively through cardiac catheterization. PVR is used as a PAH disease indicator while dp/dt max and Ees are informative variables that help physicians understand chamber stiffness. This finding holds significance by potentially mitigating the invasiveness associated with clinical PAH diagnosis. However, further research is required to unravel the underlying mechanisms and explain the medical relevance of these relationships.

Recent studies of PAH have found success leveraging SVM to perform gene expression categorization and screening for early-stage PAH diagnosis [4]. While this ML-oriented study highlights key correlations and defining parameters within our animal dataset, future improvements are required. These include acquiring and expanding our dataset to eliminate imputations and allowing the implementation of multilayer perceptron (MLP) neural networks that provide flexibility of capturing complex relationships through advanced feature learning and nonlinear transformations.

ACKNOWLEDGEMENTS

Funding provided by NHLBI 1R01HL155945-01 and NSF CAREER Award 2046259.

REFERENCES

[1] M. M. Hoeper *et al.*, *J Heart Lung Transplantation*, 39(12), 2020.
[2] P. Rambarat *et al.*, *PLoS One*, 18(8), 2023.
[3] ED Kwan et al., Am J Physiol Heart Circ Physiol, 321(4), 2021.
[4] Z. Shang *et al.*, *Front Genet*, 2021.