Echocardiographic Artificial Intelligence for Pulmonary Hypertension Classification

Yukina Hirata, PhD[1], Takumasa Tsuji, MS[2], Jun'ichi Kotoku, PhD[2], Masataka Sata, MD, PhD[3], Kenya Kusunose, MD, PhD[3,4]

[1]Ultrasound Examination Center, Tokushima University Hospital, Tokushima, Japan
[2]Department of Radiological Technology, Graduate School of Medical Care and Technology, Teikyo University, Tokyo, Japan
[3]Department of Cardiovascular Medicine, Tokushima University Hospital, Tokushima, Japan
[4]Department of Cardiovascular Medicine, Nephrology, and Neurology, Graduate School of Medicine, University of the Ryukyus, Okinawa, Japan.

Address for Correspondence:
Kenya Kusunose, MD, PhD
Department of Cardiovascular Medicine, Nephrology, and Neurology, Graduate School of Medicine, University of the Ryukyus, Okinawa, Japan.
207 Uehara, Nishihara Town, Okinawa, Japan
TEL: 81- 98-895-1150, FAX: 81- 98-895-1416
E-mail: echo.cardio@gmail.com
Twitter: @Ken_Cardiology

Abstract

Objective: The classification of pulmonary hypertension (PH) is crucial for determining the appropriate therapeutic strategy. We investigated whether machine learning (ML) algorithms may assist in echocardiographic PH prediction, where current guidelines recommend integrating several different parameters.

Methods: We obtained physical and echocardiographic data from 885 patients who underwent right heart catheterization (RHC). Patients were classified into three groups: non-PH, pre-capillary PH, and post-capillary PH, based on values obtained from RHC. Utilizing 24 parameters, we created predictive models employing four different classifiers and selected the one with the highest area under the curve (AUC). We then calculated the macro-average classification accuracy for PH on the derivation cohort (n=720) and prospective validation dataset (n=165), comparing the results with guideline-based echocardiographic assessment obtained from each cohort.

Results: Logistic regression with elastic net regularization had the highest classification accuracy, with AUCs of 0.789, 0.766, and 0.742 for normal, pre-capillary PH, and post-capillary PH, respectively. The ML model demonstrated significantly better predictive accuracy than the guideline-based echocardiographic assessment in the derivation cohort (59.4% vs. 51.6%, $p<0.01$). In the independent validation dataset, the ML model's accuracy was comparable to the guideline-based PH classification (59.4% vs. 57.8%, $p=0.638$).

Conclusions: This preliminary study suggests promising potential for our ML model in predicting echocardiographic PH. Further research and validation are needed to fully assess its clinical utility in PH diagnosis and treatment decision-making.

Keywords: Pulmonary hypertension; left heart disease; classification; echocardiography; artificial intelligence; machine learning.

WHAT IS ALREADY KNOWN ON THIS TOPIC：

・Pulmonary hypertension (PH) can be categorized into post-capillary and pre-capillary based on the underlying pathophysiology. Echocardiography is a non-invasive tool used for PH evaluation, but its interpretation for classification can be subjective and complex.

WHAT THIS STUDY ADDS：

・Utilizing a machine learning (ML) based PH classification model with 24 parameters, accurate predictions of non-PH, pre-capillary PH, and post-capillary PH were achieved, providing valuable insights into ML's potential for PH classification.

・The importance of echocardiographic indices and physical characteristics in accurate PH classification was highlighted, with LAVi and diastolic Doppler values significantly influencing post-capillary PH prediction, while other physical predictors played a crucial role in predicting the remaining two groups.

・ML model accuracy was equal or superior to a model based on TRV and diastolic grading according to ASE/EACVI guidelines.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

・The ML model demonstrates a performance that is comparable to conventional guideline-based methods in classifying subjects into likely pre- vs post-capillary PH, which could aid further diagnostic assessment.

Introduction

Pulmonary hypertension (PH) comprises two main types: post-capillary PH, linked to left heart diseases like heart failure with preserved ejection fraction (HFpEF), and pre-capillary PH, arising from pulmonary vascular remodeling. Accurate differentiation between these two types is crucial for appropriate therapeutic interventions. While PH due to left heart disease is the predominant form, characterized by hemodynamic patterns of post-capillary hypertension, its accurate diagnosis can sometimes pose challenges. This complexity is heightened by the fact that PH is present in 50-80% of HFpEF patients [1]. Although right heart catheterization (RHC) is the gold standard, its invasiveness limits routine use. Echocardiography, a non-invasive method, plays a vital role in early detection and differentiation. While several echocardiographic markers are important for assessing PH likelihood, their interpretation can be complex and subjective, relying on expertise across medical institutions.

Recent advancements in artificial intelligence (AI) in medical diagnostics have shown promise, including predicting PH or left atrial pressure (LAP) from ECG and chest X-ray[2-5]. Moreover, within the field of echocardiography, some studies reported the effective use of machine learning (ML) for the echocardiographic assessment of diastolic function [6, 7]. Omar et al. [8] reported that ML, using echocardiographic and strain analysis, predicted elevated pulmonary arterial wedge pressure (PAWP) in patients with suspected HF. Given these AI advancements, we hypothesized that an ML model, guided by echocardiographic data, could better distinguish between PH types. Our study aims to develop and validate an ML model for categorizing patients into three groups: non-PH, pre-capillary PH, and post-capillary PH.

Methods

Study population. We conducted a post-hoc analysis of two independent cohorts (**Supplemental Figure 1**). The derivation cohort comprised 1,069 patients who underwent both right heart catheterization (RHC) and echocardiography between October 2009 and February 2020. Patients were excluded for missing echocardiographic parameters (atrial fibrillation n=215, atrial flutter n=20, tachycardia n=10, complete atrioventricular block n=7, lack of tricuspid regurgitation n=55). Additionally, those with a >30-day interval between echocardiography and RHC were excluded (n=42). After applying these exclusions, 720 patients were retained in the derivation cohort, divided into three groups: non-PH, pre-capillary PH, and post-capillary PH. The pre-capillary PH group was defined as having an elevated mean pulmonary artery pressure (PAP) (>20 mmHg) and a normal range PAWP (≤15 mmHg), while the post-capillary PH group had an elevated PAWP (>15 mmHg). To ensure generalizability, a separate validation cohort (n=165) was enrolled between March 2020 and February 2023. The study protocol was approved by the Institutional Review Board of Tokushima University Hospital (no. 3217-3) and adhered to the guidelines of the Declaration of Helsinki. As data were de-identified, the requirement for informed consent was waived.

Right heart catheterization. The RHC was performed using a Swan-Ganz catheter. Patients were measured in a supine position at end-expiration. The following hemodynamic parameters were recorded: mean PAWP, mean PAP, mean right atrial pressure, and cardiac output (CO). CO was measured using the assumed Fick equation, chosen due to its reduced susceptibility to influences like tricuspid regurgitation, especially relevant for our patients with suspected PH. The pulmonary vascular resistance was defined as: *(mean PAP - mean PAWP)/CO*. Diagnosis of PH was based on these measurements, following the latest World Symposium standards [9].

Data acquisition. The median time from echocardiography to RHC was 3 days (interquartile range, 1-11). Height, weight, blood pressure and heart rate (HR) of all patients were recorded at the time of the echocardiography. The echocardiography was performed using commercially available ultrasound machines. Echocardiographic evaluations were retrospectively reclassified in alignment with the American Society of Echocardiography (ASE) and European Association of Cardiovascular Imaging (EACVI) 2015 guidelines [10]. The left ventricular (LV) and left atrial (LA) volumes were measured using the biplane method of disks on two-dimensional images. These volumes were then used to calculate the LA volume index (LAVi) and the LV ejection fraction (EF). Stroke volume (SV) was calculated from the cross-sectional area of the LV outflow tract and the time velocity integral. We used pulsed-wave Doppler to measure early (E) and late (A) diastolic flow velocities, and calculate the E/A ratio. We also recorded the pulmonary venous flow, including peak systolic (S) and diastolic (D) flow velocities, and calculated their ratio (S/D). We also performed pulsed-wave tissue Doppler imaging (TDI) at the lateral mitral annulus to measure early (e') and late (a') diastolic velocities, and calculated the E/e' ratio. Tricuspid regurgitation velocity (TRV) was measured using continuous-wave Doppler. Based on the 2016 ASE recommendation [11], DD grade II or higher was used to identify post-capillary PH (echo-post-capillary PH). In cases where the parameters were non-congruent, the DD grade was based on the highest characteristic parameter with equal weighting assumption. The echocardiographic probability of PH was determined using the 2016 ESC guidelines [12]. Patients with a TRV > 2.8 m/s and a DD grade of 0 or 1 were considered at high risk for pre-capillary PH (echo-pre-capillary PH).

Model training and testing. The study flow diagram is shown in **Figure 1**. For PH classification, we utilized 24 clinical variables (age, sex, height, weight, body mass index, HR, systolic BP,

diastolic BP, LVEDV, LVESV, LVEF, LVMi, LAVi, SVi, E, A, Dct, E/A, TDI-e', TDI-a', E/e', S/D, TRV, IVC), all standardized to have a mean of 0 and a standard deviation of 1. We divided our derivation dataset of 720 cases into a training and validation group with 648 cases (90%) and a test group with 72 cases (10%). To address imbalanced class distributions, we performed a stratified 10-fold cross-validation, ensuring each subset accurately represented the entire dataset. We used four classifiers for PH classification: logistic regression with elastic net regularization for its effectiveness in feature selection and handling multicollinearity; the Support Vector Machine (SVM) for its kernel method adaptability [13]; the Random Forest for its clear feature importance ranking [14]; and XGBoost [15], known for its efficiency and ability to fine-tune parameters. Our selection was based on their capability to effectively interpret feature importance. Model performances were evaluated by averaging the AUROC values obtained from the ten-fold cross-validation process. We then ranked the models using the Wilcoxon signed rank test, and the top model was further tested on new data to check its effectiveness and ability to generalize.

In the validation cohort, we input each patient's data into the models. For each PH class, we obtained ten predicted probabilities from these models. We then calculated the average probability for each PH class and assigned each patient to the class with the highest average probability.

Statistical analysis. Continuous data were presented as mean ± standard deviation or median (25th–75th percentile range), and categorical data as numbers and percentages. Non-normally distributed and categorical data were analyzed using the Kruskal-Wallis test and Pearson's chi-squared test, respectively. To assess the model's performance, we applied SHapley Additive exPlanations (SHAP) analysis after classifying PH using ML models [16]. This helped us see how

each factor affected our model's decisions, making it clearer and more interpretable. In addition to SHAP analysis, to evaluate the performance of our models, we employed key metrics such as accuracy, confusion matrix, and Macro Average Precision. The ML approach was compared with existing echocardiographic guidelines [11-14]. Statistical analyses were conducted using SPSS 21.0 (SPSS, Chicago, IL, USA) and MedCalc 19.5.6 (Mariakerke, Belgium), with a significance threshold set at $p < 0.05$.

Results

Results from Step 1. **Table 1** presents the outcomes of our ML models using different classifiers. We based our model selection on the AUROC values, which were computed as the average for test sets in each fold from the earlier-described cross-validation. Among all the classifiers, the logistic regression with elastic net regularization yielded the best classification accuracy. It achieved average AUCs of 0.789, 0.766, and 0.742 for the non-PH, pre-capillary PH, and post-capillary PH groups, respectively. Given its superior performance, we opted for the logistic regression with elastic net regularization for further analyses.

Clinical background. **Table 2** shows the characteristics of the study population. The derivation cohort of 720 patients and the validation cohort of 165 patients were generally similar, but the validation group had a slightly higher HR (73±14 bpm vs. 71±14 bpm, p=0.027). Echocardiography showed that the validation cohort had lower A wave velocity, S/D ratio, and TDI-e' and TDI-a' velocities, but higher E/A ratio and LAVi. There was no significant difference in TRV between the two groups. The validation cohort had fewer cases with grade I diastolic dysfunction (48% vs. 56%, p=0.047). The distribution of RHC diagnosis differed between the cohorts, with the validation cohort having more post-capillary PH cases and the derivation cohort

having more non-PH cases. This suggests there might be a class imbalance issue between the cohorts.

Results from step 2: Feature contribution for the PH model by explainable AI. **Figure 2** show the top twenty factors that help classify PH in both the derivation and validation datasets. TRV stands out as the key factor for predicting PH in our study, as shown by its high SHAP values, emphasizing its importance in PH diagnosis. However, HR and body weight also surfaced as significant predictors. The non-PH group had lower TRV, HR, and body weight, while the pre-capillary PH group showed higher TRV and HR, with more females (**Figure 3 and Supplemental Figure 2**). For post-capillary PH, echocardiographic measures like LAVi, E/A ratio, IVC, E, and TDI-a' were particularly influential. Overall, our study highlights the importance of combining echocardiographic and physical factors for accurately classifying PH.

Confusion matrices. We evaluated the efficacy of our proposed method using the test dataset and an independent validation dataset. **Figure 4** shows the confusion matrix for the prediction results of the test dataset, with color coding based on the percentage scale. Accuracy is the ratio of correct predictions to all predictions. The macro-average classification accuracy obtained from the 10-fold cross-validation showed that the ML model had a significantly greater predictive accuracy than classification using conventional echocardiographic reference values (59.4% vs. 51.6%, $p<0.01$). In separate validation dataset, the accuracy of the ML model was comparable to the guidelines-based echocardiographic assessment in the validation cohort (59.4% vs. 57.8%, $p=0.638$).

Representative case. Case 1. A patient in their 70s with HF-preserved EF who suspected ischemic heart disease was admitted. Echocardiography initially indicated an indeterminate DD grade, and the conventional echocardiographic reference classified this case into the echo-non-

PH group. However, using the ML model, we predicted that the patient had ML-post-capillary PH. Subsequently, when RHC was performed, the patient was indeed diagnosed with post-capillary PH, which was consistent with the ML prediction **(Figure 5A)**.

Case 2: A patient in their 50s, with a medical history of scleroderma, was previously diagnosed with exercise-induced pulmonary hypertension a few years ago. At that time, they chose not to undergo a RHC. During their most recent routine examination, echocardiography showed a TRV of 3.0 m/s. Our ML model predicted them as having ML-pre-capillary PH, a result that aligned with the echocardiographic findings **(Figure 5B)**. Given their recent onset of exertional dyspnea and the findings, they opted for RHC and were later admitted for the procedure. After being admitted to the hospital, the RHC findings revealed mPAP of 22 mmHg, PAWP of 6 mmHg, and PVR of 3.4 WU, confirming their diagnosis as pre-capillary PH.

Discussion

In our study, we explored how ML can improve understanding clinical and echocardiographic data for classifying PH using measurements from RHC. This is the first study to show that ML can add valuable insights to echocardiography for PH classification.

PH and its etiology. RHC is the gold standard for assessing right heart and pulmonary artery hemodynamics, but it is invasive. Noninvasive options like echocardiography are preferable [11]. However, its accuracy in matching RHC results is debated [17]. The ASE/EACVI algorithm's accuracy in identifying elevated or normal filling pressures is limited, and guidelines recommend different parameters for specific situations like PH [18-20]. Leung et al. [17] found that classifying PH can complicate the evaluation of diastolic function, resulting in confusing connections between individual diastolic parameters and filling pressures. Similarly, Sato et al.[21] demonstrated that

using guidelines to identify moderate or severe diastolic dysfunction only correctly indicated high filling pressures 66% of the time in patients who had both echocardiography and RHC. Our echocardiographic measurements aligned consistently with previous findings. In our research, the ML model's performance was on par with guideline-based classifications in differentiating post-capillary PH and even surpassed in accurately categorizing patients into non-PH, pre-capillary, and post-capillary PH groups. This ML-aided assessment offers quantitative results autonomously, which bolsters reproducibility.

However, classifying combined pre- and post-capillary PH patients posed a distinct challenge. Of the 30 patients with this profile (mean PA pressure > 20 mmHg, PAWP > 15 mmHg, and PVR > 3 WU), designated as the "post-capillary PH group" **(Supplemental Table 1)**, the ML model tended to classify 43% as ML-pre-capillary PH. This is in stark contrast to the 16% from the non-combined group ($p<0.01$). Such tendencies arise from the intricate interplay of clinical and hemodynamic markers that characterize both PH types. For instance, diseases like HFpEF, a variant of left heart disease, mainly induce post-capillary PH due to elevated left-sided filling pressures. At times, these are accompanied by pre-capillary features, complicating the categorization. The challenge amplifies when dominant pre-capillary markers, akin to those in conditions like pulmonary arterial hypertension, mask the post-capillary signals. We recognize this hurdle and are actively working to refine our ML model to better handle these overlapping characteristics. While our results spotlight the potential of ML in clinical assessments, a larger dataset will be instrumental in honing its precision, especially for nuanced classifications of combined PH profiles.

SHAP-based findings for the classification of PH. The SHAP values are a widely accepted method in the ML community to understand the impact of variables in prediction

models. The SHAP results in our study are consistent with other clinical studies, validating the ML approach in discovering patterns and important features [22]. This highlights new opportunities for the field of cardiology [23]. Other studies have also successfully utilized SHAP analyses, identifying HF clinical subtypes and depicting LV structure related to physical activity [24, 25]. Combining clinical parameters with detailed cardiovascular imaging shows promise as a novel ML application.

Our ML model, which incorporates physical and echocardiographic parameters, demonstrated high accuracy in classifying PH, often surpassing conventional methods. This novel approach could be beneficial to clinicians during routine echocardiography. TRV, serving as an indirect metric for the right ventricular systolic pressure, is elevated in the context of PH, regardless of its origin. However, it's crucial to understand that TRV alone cannot distinguish between PH secondary to left heart disease and other causes. Our analysis indicates that the likelihood of post-capillary PH is associated with increased LAVi and E/A ratio, a relatively higher E/e', a decreased TDI-a', and larger IVC dilatation. Several echocardiographic parameters, such as LAVi, TRV, and Doppler measures like E/A, e', and E/e', have been validated for diagnosing LV diastolic dysfunction. Consistently, previous studies show that patients with raised LV filling pressures tend to have a larger LAV and lateral E/e', which is consistent with our findings [26]. While TDI-a' is not a major parameter for diagnosing LV diastolic dysfunction, it is minimally affected by preload and reflects various factors influencing diastolic pressure, such as left atrial contraction and relaxation, and LV end-diastolic pressure.

Furthermore, our findings accentuate the importance of HR and weight in this domain. Elevated HR, often observed in conditions such as PAH, can be attributed to increased right ventricular afterload. Prior research has established a connection between HR and sympathetic

activity, obesity, HFpEF, and PH [27, 28]. Weight, particularly when examined within the obesity phenotype of HFpEF, emerges as a consistent prognostic indicator [29]. Our research emphasizes the vital importance of both primary and "non-specific" clinical parameters, such as age, gender, HR, and weight, in distinguishing between PH phenotypes.

Clinical implications. Early and accurate diagnosis of PH is crucial for patient care. Our ML model, still in its early stages, shows promise in aiding the initial screening for PH through echocardiography. It has performed well in our dataset, often matching or even exceeding the accuracy of traditional methods based on TRV and diastolic grading as per ASE guidelines. Recent developments like SHAP have made such ML models more interpretable and potentially useful for clinical use. Our model, as shown in **Figure 5**, provides insights that could help clinicians in making informed decisions. However, it's important to note that using this model for direct treatment or prevention strategies is still exploratory at this stage. We need more extensive validation studies with larger and more diverse patient groups to fully understand how this ML model can be used in clinical practice for PH.

Limitations. This study has several limitations. First, our small sample size could affect the results and limit the strength of our conclusions, indicating the need for larger future studies. Second, we only included patients from our center who were referred for RHC, which might not represent the general population. Third, patients with atrial fibrillation were excluded from this study due to the impact on important diastolic parameters such as the E/A ratio. Fourth, our study addressed various post-capillary PH causes (e.g., severe mitral regurgitation as highlighted in **Supplemental Table 2**). However, the main focus was distinguishing between pre-capillary and post-capillary PH using echocardiography and ML. The model might not fully capture the nuances of different post-capillary PH causes. Fifth, the echo-classification approach used in our

study based on the DD guideline's recommendation lacks formal validation, as it involved selecting the highest number of characteristic parameters when parameters were non-congruent. Finally, our study did not incorporate several echocardiographic right ventricle parameters recommended by the 2022 ESC/ERS PH Guidelines[30], such as TAPSE, RV-FAC, RV free-wall strain, and RVEF from 3D echocardiography due to non-routine clinical data collection. We aim to include these parameters prospectively to further enhance the system.

Conclusions: Our ML model shows promising potential for predicting echocardiographic PH. Further research and validation are necessary to assess its clinical utility in PH diagnosis and treatment decision-making.

Disclosures: None.

Acknowledgements:

Contributions: Design of the study: KK. Conduct of the study and data acquisition: YH. Data analysis and interpretation: TT and JK. Drafting the manuscript: HY. Reviewing the manuscript and providing input: all authors. Final approval: all authors.

Funding: This research was supported by Japan Society for the Promotion of Science Kakenhi Grants (Number 21K12706 to Y. Hirata and 23K07509 to K. Kusunose) and AMED under Grant Number JP22uk1024007 (to K. Kusunose). The funding source had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Data availability statement: Data are available upon reasonable request.

Table1. AUROC Values Evaluating Model Performance

	Elastic Net	SVM	XgBoost	Random Forest
Non-PH	0.789	0.778	0.781	0.779
Pre-capillary PH	0.766	0.758	0.753	0.751
Post-capillary PH	0.742	0.736	0.731	0.711

Results of the ML Predictive Model: The predictors were ranked according to their AUROC values, averaged over cross-validation experiments. Abbreviations: AUROC, area under receiver operating characteristics curve; SVM, support vector machine.

Table 2. Baseline characteristics of the study population.

	Derivation cohort (n=720)	Validation cohort (n=165)	P value
Age, yrs	67±14	68±15	0.206
Male, n (%)	385 (53)	91 (55)	0.348
Height, m	1.58±0.9	159±10	0.131
Weight, kg	59±13	60±15	0.129
Body surface area	1.59±0.20	1.61±0.22	0.100
HR, bpm	71±14	73±14	**0.027**
Systolic BP, mmHg	122±22	123±23	0.431
Diastolic BP, mmHg	69±14	70±14	0.213
Echocardiography			
LVEDV, mL	121±58	137±61	**<0.001**
LVESV, mL	63±49	78±54	**<0.001**
LVEF, %	53±16	48±16	**<0.001**
LVMi, g/m^2	116±44	105±30	**0.001**
LAVi, ml/m^2	43±17	45±16	**0.039**
E, cm/s	78±32	78±29	0.405
A, cm/s	80±31	72±29	**0.003**
Dct, cm	216±76	208±75	0.119
E/A ratio	1.20±0.87	1.34±1.07	**0.028**
TDI-e', cm/s	7.1±3.0	6.4±3.1	**0.007**
TDI-a', cm/s	8.6±3.5	7.7±3.4	**0.001**
E/e'	13.1±7.5	14.5±8.4	**0.019**
S/D ratio	1.3±0.6	1.2±0.6	**0.022**
SV, ml	58±21	60±21	**<0.001**
TRV, m/s	2.7±0.6	2.8±0.6	0.169
IVC, mm	12±4	12±4	0.164
DD normal, n (%)	38 (5)	11 (7)	0.158
DD grade I, n (%)	401 (56)	80 (48)	**0.047**
DD grade II, n (%)	127 (18)	34 (21)	0.187
DD grade III, n (%)	50 (7)	17 (10)	0.071
DD indeterminate, n (%)	104 (14)	22 (13)	0.356
Pulmonary hemodynamics			
Mean PAP, mmHg	21.2±9.2	22.9±9.2	0.055
Mean PAWP, mmHg	12.1±6.3	13.4±6.7	**0.010**
CO, l/min	4.6±1.4	4.5±1.5	0.144
PVR, WU	2.3±2.0	2.4±2.1	0.252
RHC diagnosis			
Non-PH, n (%)	381 (53)	75 (45)	**0.042**
Pre-capillary PH, n (%)	161 (22)	36 (22)	0.440

Post-capillary PH, n (%)	178 (25)	54 (33)	**0.017**

Data are presented as number of patients (percentage) and mean ± SD. Abbreviations: HR, heart rate; BP, blood pressure; LVEDV, left ventricular end-diastolic volume; LVESV, left ventricular end-systolic volume ; LVEF, left ventricular ejection fraction; LVMi, left ventricular mass index; LAVi, left atrial volume index; E, early diastolic transmitral flow velocity; A, late diastolic transmitral flow velocity; e', early diastolic mitral annular motion; S/D, ratio between the peak systolic and diastolic pulmonary vein velocities; SV, systolic volume; TRV, tricuspid regurgitation velocity; IVC, inferior vena cava, DD; diastolic dysfunction; PAP, pulmonary artery pressure; PAWP, pulmonary capillary wedge pressure; CO, cardiac output; PVR, pulmonary vascular resistance.

Figure legends:

Figure 1: Study Workflow. Step 1 - Model Building: We developed an ML model for PH classification using 24 clinical and echocardiographic indicators, and evaluated its performance against the gold-standard RHC data. The derivation dataset (n=720) was divided randomly, with 90% for training and validation, and 10% for testing. Stratified 10-fold cross-validation was used with four classifiers. Step 2 - Model Performance: We evaluated the importance of ML predictors using SHAP values and compared the overall classification accuracy of our ML model with the accuracy achieved through guideline-based echocardiographic assessment.

Figure 2: **A summary plot for the SHAP analysis.** (A) the derivation cohort, and (B) the validation cohort. A bar chart of the mean global feature importance to distinguish the three classes. The chart indicates the mean absolute SHAP value of the 20 most important features to differentiate the three classes, and for detecting non-PH vs. pre-capillary PH vs. Post-capillary PH.

Figure 3: A SHAP summary plot showing the effect of each feature to specifically detect each category in the derivation cohort. The colors indicate the importance of each feature to each category, (A) with non-PH depicted by green, (B) pre-capillary PH by blue, and (C) post-capillary PH by magenta.

Figure 4: **Confusion Matrixes of Classification Results for Each PH group by Logistic Regression.** Confusion matrices for (A) the derivation cohort and (B) the validation cohort. Precision is the ratio of true positives to all positives. The colors indicate the percentage of each prediction to each category, 0-9% depicted by azure, 10-29% by light blue, 30-49% by sky blue, 50-69% by steel blue, 50-69 % by and ≧70% by midnight blue.

Figure 5: Echocardiographic and physical findings before catheterization. (A) **A representative case 1.** When using echocardiography alone and adhering to the guidelines, the classification was determi**ne**d as " non-PH ". However, the ML model resulted in a "post-capillary PH" probability of 0.41, indicating the highest probability. According to SHAP analysis, weight, IVC, and e' showed significant contributions. After catheterization, the results revealed "post-capillary PH" with mPAP of 29 mmHg and PAWP of 20 mmHg. (B) **A representative case 2.** The ML model resulted in a "pre-capillary PH" probability of 0.65, indicating the highest probability. Experiencing new symptoms, they underwent an RHC, which diagnosed them with pre-capillary PH based on an mPAP of 22 mmHg and a PAWP of 6 mmHg.

Supplemental Figure1: Flow chart

Supplemental Figure 2: A SHAP summary plot showing the effect of each feature to specifically detect each category in the validation cohort.

References

1. Guazzi M and Labate V. Pulmonary Hypertension in Heart Failure Patients: Pathophysiology and Prognostic Implications. *Curr Heart Fail Rep*. 2016;13:281-294.
2. Sarker IH. Machine Learning: Algorithms, Real-World Applications and Research Directions. *SN Comput Sci*. 2021;2:160.
3. Kusunose K, Hirata Y, Tsuji T, *et al.* Deep learning to predict elevated pulmonary artery pressure in patients with suspected pulmonary hypertension using standard chest X ray. *Sci Rep*. 2020;10:19311.
4. Liu CM, Shih ESC, Chen JY, *et al.* Artificial Intelligence-Enabled Electrocardiogram Improves the Diagnosis and Prediction of Mortality in Patients With Pulmonary Hypertension. *JACC Asia*. 2022;2:258-270.
5. Kusunose K, Hirata Y, Yamaguchi N, *et al.* Deep Learning for Detection of Exercise-Induced Pulmonary Hypertension Using Chest X-Ray Images. *Front Cardiovasc Med*. 2022;9:891703.
6. Fletcher AJ, Lapidaire W and Leeson P. Machine Learning Augmented Echocardiography for Diastolic Function Assessment. *Front Cardiovasc Med*. 2021;8:711611.
7. Jiang R, Yeung DF, Behnami D, *et al.* A Novel Continuous Left Ventricular Diastolic Function Score Using Machine Learning. *J Am Soc Echocardiogr*. 2022.
8. Omar AMS, Narula S, Abdel Rahman MA *et al.* Precision Phenotyping in Heart Failure and Pattern Clustering of Ultrasound Data for the Assessment of Diastolic Dysfunction. *JACC Cardiovasc Imaging*. 2017;10:1291-1303.
9. Maron BA, Kovacs G, Vaidya A, *et al.* Cardiopulmonary Hemodynamics in Pulmonary Hypertension and Heart Failure: JACC Review Topic of the Week. *J Am Coll Cardiol*. 2020;76:2671-2681.
10. Lang RM, Badano LP, Mor-Avi V, *et al.* Recommendations for cardiac chamber quantification by echocardiography in adults: an update from the American Society of Echocardiography and the European Association of Cardiovascular Imaging. *J Am Soc Echocardiogr*. 2015;28:1-39 e14.
11. Nagueh SF, Smiseth OA, Appleton CP, *et al.* Recommendations for the Evaluation of Left Ventricular Diastolic Function by Echocardiography: An Update from the American Society of Echocardiography and the European Association of Cardiovascular Imaging. *J Am Soc Echocardiogr*. 2016;29:277-314.
12. Galie 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 Heart J*. 2016;37:67-119.
13. Chang CC, Lin CJ. LIBSVM: A Library for Support Vector Machines[Internet]. A valiable from:www.csie.ntu.edu.tw/.
14. Breiman L. Random Forests. Vol.45, 2001.
15. Chen T, Guestrin C. XGBoost: A scalable tree boosting system. In: Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. *Association for Computing Machinery*; 2016; 785-94.
16. Lundberg S, Lee SI. A Unified Approach to Interpreting Model Predictions. *Advances in Neural Information Processing System*. 2017;22:4766-75.

17. Leung EC, Swiston JR, AlAhmari L, *et al.* Validity of algorithm for estimating left sided filling pressures on echocardiography in a population referred for pulmonary arterial hypertension. *Pulm Circ*. 2017;7:2045893217740471.
18. Sharifov OF, Schiros CG, Aban I, *et al.* Diagnostic Accuracy of Tissue Doppler Index E/e' for Evaluating Left Ventricular Filling Pressure and Diastolic Dysfunction/Heart Failure With Preserved Ejection Fraction: A Systematic Review and Meta-Analysis. *J Am Heart Assoc*. 2016;5.
19. Nauta JF, Hummel YM, van der Meer P, *et al.* Correlation with invasive left ventricular filling pressures and prognostic relevance of the echocardiographic diastolic parameters used in the 2016 ESC heart failure guidelines and in the 2016 ASE/EACVI recommendations: a systematic review in patients with heart failure with preserved ejection fraction. *Eur J Heart Fail*. 2018;20:1303-1311.
20. Jones R, Varian F, Alabed S, *et al.* Meta-analysis of echocardiographic quantification of left ventricular filling pressure. *ESC Heart Fail*. 2021;8:566-576.
21. Sato K, Grant ADM, Negishi K, *et al.* Reliability of updated left ventricular diastolic function recommendations in predicting elevated left ventricular filling pressure and prognosis. *Am Heart J*. 2017;189:28-39.
22. Zeng X, Hu Y, Shu L, *et al.* Explainable machine-learning predictions for complications after pediatric congenital heart surgery. *Sci Rep*. 2021;11:17244.
23. Wang K, Tian J, Zheng C, *et al.* Interpretable prediction of 3-year all-cause mortality in patients with heart failure caused by coronary heart disease based on machine learning and SHAP. *Comput Biol Med*. 2021;137:104813.
24. Lu S, Chen R, Wei W, *et al.* Understanding Heart Failure Patients EHR Clinical Features via SHAP Interpretation of Tree-Based Machine Learning Model Predictions. *AMIA Annu Symp Proc*. 2021;2021:813-822.
25. Angelaki E, Marketou ME, Barmparis GD, *et al.* Detection of abnormal left ventricular geometry in patients without cardiovascular disease through machine learning: An ECG-based approach. *J Clin Hypertens (Greenwich)*. 2021;23:935-945.
26. Andersen OS, Smiseth OA, Dokainish H, *et al.* Estimating Left Ventricular Filling Pressure by Echocardiography. *J Am Coll Cardiol*. 2017;69:1937-1948.
27. Molfino A, Fiorentini A, Tubani L, *et al.* Body mass index is related to autonomic nervous system activity as measured by heart rate variability. *Eur J Clin Nutr*. 2009;63:1263-5.
28. Dachs TM, Duca F, Rettl R, *et al.* Riociguat in pulmonary hypertension and heart failure with preserved ejection fraction: the haemoDYNAMIC trial. *Eur Heart J*. 2022;43:3402-3413.
29. Borlaug BA, Kitzman DW, Davies MJ, *et al.* Semaglutide in HFpEF across obesity class and by body weight reduction: a prespecified analysis of the STEP-HFpEF trial. *Nat Med.* 2023;29: 2358-2365
30. Humbert M, Kovacs G, Hoeper MM, *et al.* 2022 ESC/ERS Guidelines for the diagnosis and treatment of pulmonary hypertension: Developed by the 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 the International Society for Heart and Lung Transplantation (ISHLT) and the European Reference Network on rare respiratory diseases (ERN-LUNG). *Eur Heart J*. 2022;43:3618-3731.