Artificial Intelligence Analysis of Risk Factors for Preterm Patent Ductus Arteriosus: Nationwide Data

Jae Yoon Na
Department of Pediatrics, Hanyang University College of Medicine
Dongkyun Kim
Department of Electrical and Electronic Engineering, Hanyang University
Amy M. Kwon
Biostatistics Core, HY Medicine-Engineering-Bio (MEB): Global Centre for Developmental Disorder, Industry-University Cooperation, Hanyang University
Jin Yong Jeon
Department of Architectural Engineering, Hanyang University
Hyuk Kim
Department of Thoracic and Cardiovascular Surgery, Hanyang University
Chang-Ryul Kim
Department of Pediatrics, Hanyang University College of Medicine
Hyun Ju Lee
Department of Pediatrics, Hanyang University College of Medicine
Joohyun Lee
Department of Electrical and Electronic Engineering, Hanyang University
Hyun-Kyung Park (neopark@hanyang.ac.kr)
Department of Pediatrics, Hanyang University College of Medicine

Research Article

Keywords: Artificial Intelligence, Analysis, Risk Factors, Preterm Patent, Ductus Arteriosus, Nationwide Data, participants

Posted Date: September 8th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-861626/v1

Abstract

Despite the many comorbidities and the high mortality rate in preterm infants with patent ductus arteriosus (PDA), therapeutic strategies vary depending on the clinical setting, and most studies of the related risk factors have been based on small sample populations. The goal was to compare the accuracy of artificial intelligence (AI) analysis with that of conventional analysis for identifying the risk factors associated with symptomatic PDA (sPDA) in very low birth weight infants.

This nationwide prospective cohort study included 8,369 infants. The participants were divided into an sPDA group and a nonPDA group, and The sPDA group was further divided into treated and untreated subgroups. A total of 47 perinatal risk factors were collected and analyzed. Multiple logistic regression as a standard analytic tool, five AI algorithms were used to identify the factors associated with sPDA. The combination of a large database of risk factors from nationwide registries and AI techniques achieved higher accuracy and better performance on the PDA prediction tasks, and the ensemble methods showed the best performances.

Introduction

Ductus arteriosus should exist during fetal periods, when circulation in the lungs and the body is normally supplied by the mother; in term infants, the ductus arteriosus becomes functionally closed by 72 hours of age[1,2]. Approximately 20 – 50% of neonates have a gestational age (GA) < 32 weeks on day 3 of life[3], and up to 60% have a GA < 29 weeks[2,4]. Patent ductus arteriosus (PDA) in preterm infants results in increased mortality and morbidities, but the criteria for symptomatic PDA (sPDA) and the methods/timing of PDA treatment remain controversial depending on the clinical setting[5–7].

AI is the ability of a computer to simulate human intelligence based on substantial amounts of data, sophisticated algorithms, and high computational power[8]. AI can be classified into several subfields, such as machine learning, deep learning, or cognitive computing[9]. AI technologies are growing in use in the fields of imaging, diagnosis, therapy selection, risk prediction, disease stratification, and precision medicine[10]. This study is the first report to analyze the perinatal risk factors of preterm sPDA cases registered in a nationwide cohort database and to suggest the feasibility of AI in newborn screening for this disease.

The purpose of this study was to investigate the perinatal risk factors leading to sPDA and sPDA treatments for VLBW infants in a nationwide cohort registry and to compare the accuracy of AI analysis with that of conventional analysis. This study may support the idea that an integrated combination of AI and conventional analysis can synergistically aid in clinical risk prediction and therapy selection in medicine.

Methods

Study Design

Patients and Data Collection

For this study, we derived data from infants registered in the Korean Neonatal Network (KNN), a national multicenter neonatal network built on a prospective web-based registry. These data were collected from patients admitted to 74 NICUs in Korea. The KNN registry was approved by the institutional review board of each participating hospital. Informed consent was obtained from the parents of each infant prior to participation in the KNN registry. All methods were performed in accordance with the relevant guidelines and regulations. This research was supported by the Korea Centers for Disease Control and Prevention (2019-ER7103-01)[22]. This study was approved by the Hanyang University Institutional Review Board (IRB No. 2013-06-025-043).

The cohort data consisted of 10,390 VLBW infants born between January 5, 2013, and November 19, 2017, weighing less than 1,500 grams. Since confirmation of sPDA was impossible, infants whose hospital stays were less than four days, infants who received PDA treatment within three days of birth and infants who had major congenital anomalies were excluded. After this exclusion, 8,369 infants from the KNN were eligible for sPDA prediction and risk factor analysis. After the group with no PDA was excluded, 2,982 patients' data remained and were used to analyze the treatment-determining factors of sPDA (Figure 3).

Definitions

The nonPDA group (nPDA) comprised infants who had nonsymptomatic PDA or spontaneously closed PDA. sPDA was defined as the presence of more than 2 of the following 5 clinical symptoms with/without echocardiographic confirmation of a large left-to-right ductal flow: (1) a systolic or continuous murmur; (2) a bounding pulse or hyperactive precordial pulsation; (3) hypotension; (4) respiratory difficulty; and (5) evidence from a chest radiograph (pulmonary congestion, cardiomegaly). According to the therapeutic strategies for PDA registered in the KNN database, we further stratified the sPDA group into the following two subgroups to compare the risk factors in the treatment and nonintervention groups that mediated sPDA closure: the treated group (sPDA_tx), comprising patients who received any treatment for sPDA, and the untreated group (sPDA_non-tx), comprising patients who received only conservative treatment or no treatment for sPDA. The term "treatment" refers to medication (indomethacin, ibuprofen, and other NSAIDs) or ligation.

Artificial Intelligence Analysis

Dataset and Preparation

The obtained cohort data included (1) 23 factors related to the prenatal environment and pregnancy; (2) 21 factors associated with delivery and the period immediately after birth; and (3) 3 factors recorded after birth in the clinical data. Except for factors with an ambiguous causal relationship with PDA, all data from the KNN were used to the greatest extent possible. Each factor was classified as a continuous, ordinal, or nominal variable. The details of each risk factor and the abbreviations used in this study are presented in Supplementary Table 1. We imputed the missing values using the medians of continuous and ordinal variables and the modes of nominal variables.

Artificial Intelligence Algorithms

Five classical AAs were selected in this study since their specific properties are appropriate for risk prediction analyses: a random forest (RF), a decision tree-based theory used to avoid overfitting; a light gradient boosting machine (L-GBM), a low-bias model formed by combining sequential weak models with a light computational algorithm; a multilayer perceptron (MLP), a feedforward artificial neural network that has excellent pattern extraction capability; a support vector machine (SVM), a model optimized by taking advantage of the kernel trick for highly complex problems in cases where linear separation is not possible; and k-nearest neighbors (k-NN), which performs classification based on the majority of data points that are nearby. Of these AAs, the RF and L-GBM are ensemble models. The AAs used in this study have been previously used to predict hypertension and cardiovascular risk[9,16]. Further detailed information about these AAs is provided in Supplementary Methods 1 and 2.

Hyperparameter Optimization

The study population was divided into a training set and a test set at a ratio of 80:20 using stratified random sampling[23]. We resolved the class imbalance between the outcomes with the synthetic minority oversampling technique (SMOTE) algorithm to prevent overfitting of the majority class[24]. We found the optimal AI hyperparameters that maximized the area under the receiver operating curve (AUC) by a grid search and 5-fold cross validation. We evaluated the performance of the AAs by the AUC and accuracy metrics, and 95% confidence intervals (CIs) were calculated using bootstrapping[25]. We implemented the AAs using Python 3.8.5 (Python Software Foundation) and a compatible package, i.e., Scikit-Learn version 0.24[26].

Shapely Additive Explanation

After training the AAs, we analyzed the associations between the risk factors and the outcomes. AI classifiers are black boxes that do not reveal their internal working processes, making it hard to understand the associations between specific factors and decisions. To give clinicians a convincing reason to trust the decisions, we used a game theory-based AI interpretation method called SHAP (Shapely Additive exPlanations)[27]. SHAP is a leading algorithm for identifying the main risk factors that drive the decisions of a model. The SHAP value of each factor was calculated as the average difference in the prediction probabilities between the combinations of risk factors in which the target risk factor was included and not included. Due to their computational nature, SHAP values can be positive or negative depending on the side to which the given risk factor pushes the model's predictions.

MLR Analysis

We also evaluated the predictive accuracies of the examined risk factors using a conventional analysis. An MLR approach was used as the reference method, and the raw data were stratified into a binomial distribution. Variables with a threshold p-value of 0.15 were selected to remain in the model according to backward selection, starting with the full model, and all regression coefficients were tested by using Wald statistics at α=0.05. The effects whose p-values were less than 0.05 were regarded as significant. The goodness of fitness of the final model was tested using Hosmer-Lemeshow's method at α=0.05. The identification of significant effects was based on SAS 9.4 (SAS, Inc., NC, USA), and prediction analyses were performed using Python 3.7.5.

Interpretation of the Correlations among the Factors

In addition to the main AI-based risk factor analysis, we analyzed the statistical correlations among all the risk factors and outcomes. A correlation matrix was calculated using formulas such as Spearman's rank, point biserial coefficients, and pi coefficients, depending on the type of variable (continuous, ordinal, nominal), across the entire study cohort (n = 8,369). The correlation results were visualized as dendrograms, heat maps and networks through a hierarchical clustering process using the Ward distance method and the Force Atlas function of the gephi 0.9.2 program[28-30].

Results

Study Population

A total of 10,390 infants were born between January 5, 2013, and November 19, 2017, and met the KNN's inclusion/exclusion criteria for the original cohort. Among them, 2,982 (35.6%) patients had sPDA, and 5,387 (64.4%) patients did not. Of the patients with sPDA, 2465 (82.7%) were treated, and 517 (17.3%) were not treated (Table 1). The same variables were collected for these two study populations.

Table 1
Demographic Characteristics of the Study Population (N = 8,369)

Characteristic	N (%)	Mean ± SD
Gestational age (weeks)		29.1 ± 2.9
< 26	1,258 (15.0)	
26 – 29	2,870 (34.3)	
30 – 33	2,381 (28.5)	
34 – 37	530 (6.3)	
≥ 37	1,330 (15.9)	
Birth weight (g)		1,105.1 ± 276.6
< 500	131 (1.6)	
500 – 999 g	2,736 (32.7)	
1000 – 1500 g	5,502 (65.7)	
Birth height (cm)		36.7 ± 3.6
Birth head circumference (cm)		26.1 ± 2.4
Male sex	4,232 (50.6)	
Multiple births (≥ 2)	2,935 (35.1)	
Cesarean section	1,798 (21.5)	
Grouping by PDA status		
Symptomatic PDA (sPDA)	2,982 (35.6)	
with any treatment[a] (sPDA_tx)	2,465 (29.5)	
without treatment (sPDA_nontx)	517 (6.2)	
Asymptomatic PDA or spontaneously closed PDA (nPDA)	5,387 (64.4)	

Abbreviations: SD, standard deviation; PDA, patent ductus arteriosus.

[a] Treatments for PDA included medications, such as indomethacin and ibuprofen, as well as ligation surgery.

Prediction Accuracy (AI versus MLR)

The sensitivity, specificity, accuracy and AUC of the multiple logistic regression (MLR) and each AI algorithm (AA) are presented in Table 2. L-GBM (AUC, 0.819 [95% CI, 0.796–0.840]) was the best at predicting sPDA, followed by RF (AUC, 0.813 [95% CI, 0.789–0.835]) and MLP (AUC, 0.812 [95% CI, 0.790–0.832]). The RF (AUC, 0.815 [95% CI, 0.765–0.863]) and L-GBM (AUC, 0.804 [95% CI, 0.758–0.849]) models produced higher AUCs than those of the other models in determining sPDA_tx (Supplementary Fig. 1).

Table 2
Performance Metrics of the Algorithms for Predicting sPDA and sPDA_tx

	sPDA				sPDA_tx			
	Accuracy	AUC	Sensitivity	Specificity	Accuracy	AUC	Sensitivity	Specificity
MLR	*0.76 (0.74–0.78)*	*0.81 (0.79–0.83)*	*0.85 (0.83–0.87)*	*0.60 (0.58–0.62)*	*0.85 (0.82–0.87)*	*0.78 (0.74–0.81)*	*0.85 (0.28–0.32)*	*0.98 (0.97–0.99)*
RF	0.76 (0.74–0.78)	0.81 (0.79–0.84)	0.64 (0.60–0.68)	0.83 (0.81–0.85)	**0.85 (0.82–0.88)**	**0.82 (0.77–0.86)**	**0.97 (0.96–0.99)**	0.36 (0.28–0.45)
L-GBM	**0.77 (0.75–0.79)**	**0.82 (0.80–0.84)**	0.65 (0.61–0.69)	**0.84 (0.81–0.86)**	0.85 (0.82–0.87)	0.80 (0.76–0.85)	0.93 (0.90–0.95)	0.34 (0.26–0.41)
MLP	0.75 (0.73–0.77)	0.81 (0.79–0.83)	0.75 (0.72–0.78)	0.74 (0.72–0.77)	0.77 (0.73–0.80)	0.72 (0.66–0.77)	0.83 (0.80–0.86)	0.52 (0.44–0.61)
SVM	0.75 (0.73–0.78)	0.81 (0.79–0.84)	**0.76 (0.73–0.79)**	0.75 (0.73–0.78)	0.77 (0.74–0.81)	0.77 (0.72–0.82)	0.82 (0.79–0.86)	**0.57 (0.48–0.66)**
k-NN	0.66 (0.64–0.69)	0.74 (0.72–0.77)	0.73 (0.70–0.76)	0.63 (0.60–0.66)	0.67 (0.63–0.71)	0.67 (0.61–0.72)	0.71 (0.67–0.75)	0.49 (0.40–0.58)

Abbreviations: PDA, patent ductus arteriosus; CI, confidence interval; AUC, area under the receiver operating characteristic curve; MLR, multilinear regression; RF, random forest; L-GBM, light gradient boosting machine; MLP, multilayer perceptron; SVM, support vector machine; k-NN, k-nearest neighbors.

[a] The performances are presented as the mean values (95% CI).

[b] The numbers in italics show the factor analysis with MLR as the standard reference method.

The underlined values denote the highest accuracy and AUC results.

Variable Rankings

The important factors for the AI classifiers were ranked by the average absolute SHAP values, and Fig. 1 lists up to 10 important risk factors for each model. We considered factors with SHAP values greater than 0.20 as important factors and present those risk factors for each model in Table 3. These procedures were performed separately for sPDA and sPDA_tx. The full rankings of the variables in the AI analysis are shown in Supplementary Table 2.

Table 3
Top Significant Variables for sPDA and sPDA_tx Prediction

	Standard	Artificial intelligence algorithms				
	MLR[b]	**RF**	**L-GBM**	**MLP**	**SVM**	**k-NN**
sPDA vs nPDA	*SEPS*	*I_VENT*	*I_VENT*	GA	GA	GA
	pH	GA	GA	*I_VENT*	*I_VENT*	WT
	I_VENT	SEPS	SEPS	WT	WT	*I_VENT*
	BPL	WT	WT	SEPS	*SFT (n)*	*SFT (n)*
	POLY	*SFT (n)*	*SFT (n)*	*SFT (n)*	SEPS	HT
	SFT (n)	HT	HT	PROM	HT	ANS
	EPI_R			PARITY	PROM	HC
	GA			HT		
	PROM			*HC*		
	NI_VENT			*GRAV*		
sPDA_tx vs sPDA_nontx	pH	SEPS	SEPS	SEPS	SEPS	*O2*
	SEPS	*O2*	*O2*	*PROM*	*O2*	SFT (n)
	CPR_R	*O2_R*	*O2_R*	*O2*	*O2_R*	ANS
	O2_R	*NI_VENT*	*TEMP*	*TEMP*	*NI_VENT*	*TEMP*
	NI_VENT			*O2_R*	*TEMP*	MULTI
	BPL			GRAV	GRAV	GRAV
	TEMP			MULTI	F_EDU	SEPS
	RDS			HC	*PARITY*	MULTI (th)
	O2			MULTI (th)	SFT (n)	*WT*

Abbreviations: PDA, patent ductus arteriosus; sPDA, symptomatic PDA; sPDA_tx, sPDA with any treatment; RF, random forest; L-GBM, light gradient boosting machine; MLP, multilayer perceptron; SVM, support vector machine; k-NN, k-nearest neighbors. The abbreviations for all factors are shown in Supplementary Table 1.

[a] Feature importance describes how relevant a factor is to the model's predictions. In MLR, the feature importance values were selected according to a p-value of 0.05 during the testing procedure. These are listed in descending order as the absolute values of the coefficients for the MLR and as the average absolute SHAP values for the AAs.

[b] The gray column shows the factor analysis with MLR as the standard reference method.

Italics: Positive associations between the selected factors and sPDA or sPDA_tx.

	OLIGO	5_AS	*WT*	5_AS

Abbreviations: PDA, patent ductus arteriosus; sPDA, symptomatic PDA; sPDA_tx, sPDA with any treatment; RF, random forest; L-GBM, light gradient boosting machine; MLP, multilayer perceptron; SVM, support vector machine; k-NN, k-nearest neighbors. The abbreviations for all factors are shown in Supplementary Table 1.

[a] Feature importance describes how relevant a factor is to the model's predictions. In MLR, the feature importance values were selected according to a p-value of 0.05 during the testing procedure. These are listed in descending order as the absolute values of the coefficients for the MLR and as the average absolute SHAP values for the AAs.

[b] The gray column shows the factor analysis with MLR as the standard reference method.

Italics: Positive associations between the selected factors and sPDA or sPDA_tx.

Positive/Negative Correlation Analysis

The summary plot of SHAP in Supplementary Fig. 2 shows the quantitative contributions of the top 10 factors for sPDA/nPDA and sPDA_tx/sPDA_nontx in the conducted AI analysis. For example, we found that invasive mechanical ventilator treatment and the number of administered surfactants were positively associated with sPDA, and gestational age, sepsis, birth weight and birth height were negatively correlated. For sPDA_tx, supplemental oxygen, the need for oxygen supplementation at birth, and noninvasive mechanical ventilator treatment were positively associated with sepsis. Antenatal steroid use was negatively correlated.

Relationships among Risk Factors

Hierarchical clustering was used to cluster highly correlated factors in a dendrogram (Fig. 2a). sPDA was clustered with gestational period, birth height, birth weight, and birth head circumference. sPDA_tx was clustered with sepsis and fungal infections, and its cluster was closest to the cluster consisting of noninvasive mechanical ventilation, oxygen inhalation, birth temperature, etc.

According to the heat map (Fig. 2b), sPDA was negatively correlated with the gestational period, birth height, birth weight, Apgar score, and head circumference. In contrast, it was positively correlated with surfactant administration, positive airway pressure, and endotracheal intubation. sPDA_tx showed a negative correlation with sepsis and fungal infection and a positive correlation with noninvasive mechanical ventilation and oxygen inhalation.

As Fig. 2c shows, the sPDA was relatively close to the Apgar score, physical measurement, resuscitation, and ventilator treatment factors. This means that regardless of whether those factors were positively or negatively correlated, they were highly correlated. On the other hand, parental factors were far from sPDA, indicating that they had no correlation with sPDA. sPDA_tx was located near sepsis and oxygen supplementation. The factors that were highly correlated with sPDA or sPDA_tx are also shown as important factors in Table 3; therefore, the important factors derived from the AAs were somewhat consistent.

Discussion

The analysis of risk factors for symptomatic PDA and the determination of PDA treatment in VLBW infants using AI showed higher accuracy and better performance than the conventional analysis. The ensemble model showed a better prediction accuracy and AUC than the other methods when the performances of the models were evaluated. Ultimately, conventional analysis and AI analysis were incorporated to create a new diagram containing the relationships between each factor and sPDA to allow medical staff to intuitively apply these results in actual clinical practice.

According to a relatively recent large-scale study conducted in another country, respiratory distress syndrome (RDS), birth weight, female sex, gestational age, and 5-minute Apgar scores were suggested as risk factors for sPDA in preterm infants[11]. In a study that analyzed 18 factors for hemodynamically significant PDA in preterm infants within 22-29 weeks of gestational age, lower gestational age, pregnancy-induced hypertension (PIH), and surfactant use were analyzed as risk factors[12]. In this study, a very low gestational age, low birth weight and height, and the number of administered surfactants showed close correlations with sPDA in general. However, the presence of RDS and PIH did not significantly affect the prediction of sPDA. Instead, in the case of invasive ventilator care, a clear prediction of sPDA was shown. In the case of sepsis, the opposite result was obtained with MLR and other AAs. For the national data used in this study, the definition of sepsis was limited to patients who had positive blood cultures or received more than 5 days of systemic antibiotic treatment. Therefore, the definition of neonatal sepsis is not clear[13], and the inability to include culture-negative sepsis in particular led to the creation of statistical bias.

For the prediction of sPDA treatment, AI showed very high accuracy and good performance. Since there are no existing studies that determine the presence or absence of treatment for sPDA, it is impossible to know exactly which model selects the most accurate factors. In general, it was determined that there is a relatively high probability of that PDA will be treated when there is no sepsis, when supplemental oxygen is provided during the hospitalization (O2) or is needed at birth (O2_R) and when noninvasive ventilator care is required (NI_VENT). When the AAs and MLR were examined in detail, some differences were found; these were thought to be due to differences in the strategies used at various hospitals, even in cases in which the sPDA situations are the same and when the time point of sPDA and the time point at which treatment is started are different.

In recent years, AI has been applied and used in a wide variety of fields beyond simple engineering domains, and advances in machine learning have begun to affect real-world decisions in many areas, including politics, economics, finance, and medicine[14,15]. The applications of AI in health care include image analysis, treatment, the diagnosis and prognosis of diseases, health care, the improvement of medical administration and management systems, and drug development, and in some studies, AI has shown sufficient or rather high disease risk prediction ability compared to existing models[16,17].

To our knowledge, this is the first study to use AAs to predict sPDA and sPDA_tx and to analyze the main risk factors for sPDA using large-scale cohort data consisting of only electronic records and structured factors (excluding images). The proposed AI classifiers can classify patients well, even when there are nonlinear relationships in the data. This nonlinear characteristic makes it difficult to interpret the prediction processes of AI classifiers. However, by introducing a game-theoretical contribution-based explanation algorithm

(SHAP), we were able to identify the main factors. The fact that the AI models' performances exceeded AUCs of 0.8 and that the main factors were identified using a mathematically fair explanation method supports the study's validity. In addition, the main factors derived from AI and the factors that were statistically highly correlated with the outcomes coincided, thereby enhancing the consistency of the AI and the correlation analysis results.

As Table 3 shows, too many factors were considered by the MLP, the SVM, and k-NN, indicating that these models overfitted trivial factors and had lower performances. On the other hand, the tree-based ensemble models achieved the highest performances, as they are designed to overcome overfitting. Ensemble learning has been demonstrated to be a solution for constructing balanced datasets to enhance prediction performance.

However, AI analysis still has some limitations, such as representation, accuracy, and homogeneity, that occur during the data collection process[18], and the nature of self-extracting data from large datasets makes it difficult to determine how an AI method produces results and why errors occur[19,20]. Overreliance on AI models when making decisions or analyzing images may lead to automation bias, and it is difficult to analyze the basis of a given judgment[21]. Furthermore, since a SHAP value is a measure of the corresponding factor's contribution to the model result, it is impossible to predict the amount of change induced in a model's prediction based on a change in factor value. Moreover, the data collected by the KNN are not focused on PDA, so limited factors are included. The lack of information, including vital signs, may reduce the performance of AI. In the end, AI is still indispensable for use by medical staff who treat patients directly in clinical practice.

To overcome the abovementioned limitations, this study attempted to enhance objectivity by conducting integrated analysis of MLR, which has been widely used, and AAs. To consider the effects of multicollinearity, Figure 2c was used to understand the interrelationships among the factors so that the factors that were most strongly correlated with PDA could recognize the influence of one another more readily than they recognize their influence on PDA. Although this study used existing AAs, in future studies, we will try to analyze risk factors by developing a new model explanation method that detects the amounts of changes in risk factors, analyzes the relationships between treatment methods and the long-term prognosis according to the timing of a given treatment and proposes the best treatment policy.

This AI analysis using a nationwide cohort registry is the first study of VLBW infants in the neonatal intensive care unit (NICU). We were able to evaluate risk factor variables associated with and potentially causally linked to sPDA and sPDA_tx and to show that the ensemble models (RF and L-GBM) were the best among the examined AAs at predicting specific disease development trends, yielding higher accuracy than that of an established risk prediction approach. The use of these readily available online AAs underlined their applicability as an auxiliary means of risk prediction and therapy selection.

Declarations

Data Availability

Data available: Yes

Data types: Deidentified participant data

How to access data: neopark@hanyang.ac.kr; joohyunlee@hanyang.ac.kr

When available: Upon publication

Additional Information

Who can access the data: Researchers whose proposed use of the data has been approved

Types of analyses: For any purpose

Mechanisms of data availability: After the approval of a proposal

Acknowledgements

Funding

This research was supported by the Bio & Medical Technology Development Program of the National Research Foundation of Korea (NRF) funded by the Korean government (MSIT) [grant number NRF-2019M3E5D1A01069363].

Author Contributions

Na and Park had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Study concept and design: Na, Kwon, J Lee, Park

Acquisition, analysis, or interpretation of data: Na, D Kim, Kwon, HJ Lee, J Lee

Drafting of the manuscript: Na, D Kim, Kwon, J Lee, Park

Critical revision of the manuscript for important intellectual content: All authors

Statistical analysis: AM Kwon, D Kim, J Lee

Obtained funding; JY Jeon, J Lee, Park

Administrative, technical, or material support: Na, D Kim, JY Jeon, HJ Lee

Study supervision: H Kim, CR Kim, J Lee, H-K Park

Competing Interests

The authors declare no competing interests.

References

1. Hoffman, J. I. E. & Kaplan, S. The incidence of congenital heart disease. *J. Am. Coll. Cardiol.***39,** 1890-1900 (2002).

2. Benitz, W. E. Patent ductus arteriosus in preterm infants. *Pediatrics***137,** e20153730 (2016).

3. Sellmer, A. et al. Morbidity and mortality in preterm neonates with patent ductus arteriosus on day 3. *Arch. Dis. Child. Fetal. Neonatal. Ed.***98,** F505-F510 (2013).

4. Weisz, D. E. & McNamara, P. J. Patent ductus arteriosus ligation and adverse outcomes: causality or bias? *J. Clin. Neonatol.***3,** 67-75 (2014).

5. Bose, C. L. & Laughon, M. M. Patent ductus arteriosus: lack of evidence for common treatments. *Arch. Dis. Child. Fetal. Neonatal. Ed.***92,** F498-F502 (2007).

6. Sung, S. I., Lee, M. H., Ahn, S. Y., Chang, Y. S. & Park, W. S. Effect of nonintervention vs oral ibuprofen in patent ductus arteriosus in preterm infants: a randomized clinical trial. *JAMA Pediatr.***174,** 755-763 (2020).

7. Clyman, R. I. et al. PDA-TOLERATE trial: an exploratory randomized controlled trial of treatment of moderate-to-large patent ductus arteriosus at 1 week of age. *J. Pediatr.***205,** 41-48.e6 (2019).

8. Obermeyer, Z. & Emanuel, E. J. Predicting the future—big data, machine learning, and clinical medicine. *N. Engl. J. Med.***375,** 1216 (2016).

9. Krittanawong, C. et al. Future direction for using artificial intelligence to predict and manage hypertension. *Curr. Hypertens. Rep.***20,** 75 (2018).

10. Kalfa, D., Agrawal, S., Goldshtrom, N., LaPar, D. & Bacha, E. Wireless monitoring and artificial intelligence: a bright future in cardiothoracic surgery. *J. Thorac. Cardiovasc. Surg.***160,** 809-812 (2020).

11. Pourarian, S., Farahbakhsh, N., Sharma, D., Cheriki, S. & Bijanzadeh, F. Prevalence and risk factors associated with the patency of ductus arteriosus in premature neonates: a prospective observational study from Iran. *J. Matern.-Fetal Neonatal Med.***30,** 1460-1464 (2017).

12. Lee, J. A., Sohn, J. A., Oh, S. & Choi, B. M. Perinatal risk factors of symptomatic preterm patent ductus arteriosus and secondary ligation. *J. Pediatr. Neonatol.***61,** 439-446 (2020).

13. Wynn, J. L. & Polin, R. A. Progress in the management of neonatal sepsis: the importance of a consensus definition. *J. Pediatr. Res.***83,** 13-15 (2018).

14. Petrasic, K., Saul, B., Greig, J., Bornfreund, M. & Lamberth, K. Algorithms and bias: what lenders need to know. *White Case* (2017).

15. Barry-Jester, A. M., Casselman, B. & Goldstein, D. in *The Marshall Project* (2015).

16. Alaa, A. M., Bolton, T., Di Angelantonio, E., Rudd, J. H. F. & van der Schaar, M. Cardiovascular disease risk prediction using automated machine learning: a prospective study of 423,604 UK Biobank participants. *PLoS One***14,** e0213653 (2019).

17. Safavi, K. C. et al. Development and validation of a machine learning model to aid discharge processes for inpatient surgical care. *JAMA Netw. Open***2,** e1917221 (2019).

18. Im, S. H., Jung, Y. & Kim, S. H. Current status and future direction of biodegradable metallic and polymeric vascular scaffolds for next-generation stents. *Acta Biomater.***60,** 3-22 (2017).

19. Rajkomar, A., Dean, J. & Kohane, I. Machine learning in medicine. *N. Engl. J. Med.***380,** 1347-1358 (2019).

20. Murdoch, W. J., Singh, C., Kumbier, K., Abbasi-Asl, R. & Yu, B. Interpretable machine learning: definitions, methods, and applications. *arXiv:1901.04592* (2019).

21. Weng, S. F., Reps, J., Kai, J., Garibaldi, J. M. & Qureshi, N. Can machine-learning improve cardiovascular risk prediction using routine clinical data? *PLoS One***12,** e0174944 (2017).

22. Chang, Y. S., Park, H. Y. & Park, W. S. The Korean neonatal network: an overview. *J. Korean Med. Sci.***30,** S3-S11 (2015).

23. Singh, R. & Mangat, N. S. *Elements of Survey Sampling* (Springer Science & Business Media, 2013).

24. Chawla, N. V., Bowyer, K. W., Hall, L. O. & Kegelmeyer, W. P. SMOTE: synthetic minority over-sampling technique. *J. Artif. Intell. Res.***16,** 321-357 (2002).

25. Cortes, C. & Mohri, M. Confidence intervals for the area under the ROC curve. *Adv. Neural Inf. Process. Syst.***17,** 305-312 (2005).

26. Pedregosa, F. et al. Scikit-learn: machine learning in python. *J. Mach. Learn. Res.***12,** 2825-2830 (2011).

27. Lundberg, S. & Lee, S. I. A unified approach to interpreting model predictions. *arXiv:1705.07874* (2017).

28. Johnson, S. C. Hierarchical clustering schemes. *Psychometrika***32,** 241-254 (1967).

29. Jacomy, M., Venturini, T., Heymann, S. & Bastian, M. ForceAtlas2, a continuous graph layout algorithm for handy network visualization designed for the Gephi software. *PLoS One***9,** e98679 (2014).

30. Ward, J. H. Hierarchical grouping to optimize an objective function. *J. Am. Stat. Assoc.***58,** 236-244 (1963).

Figures

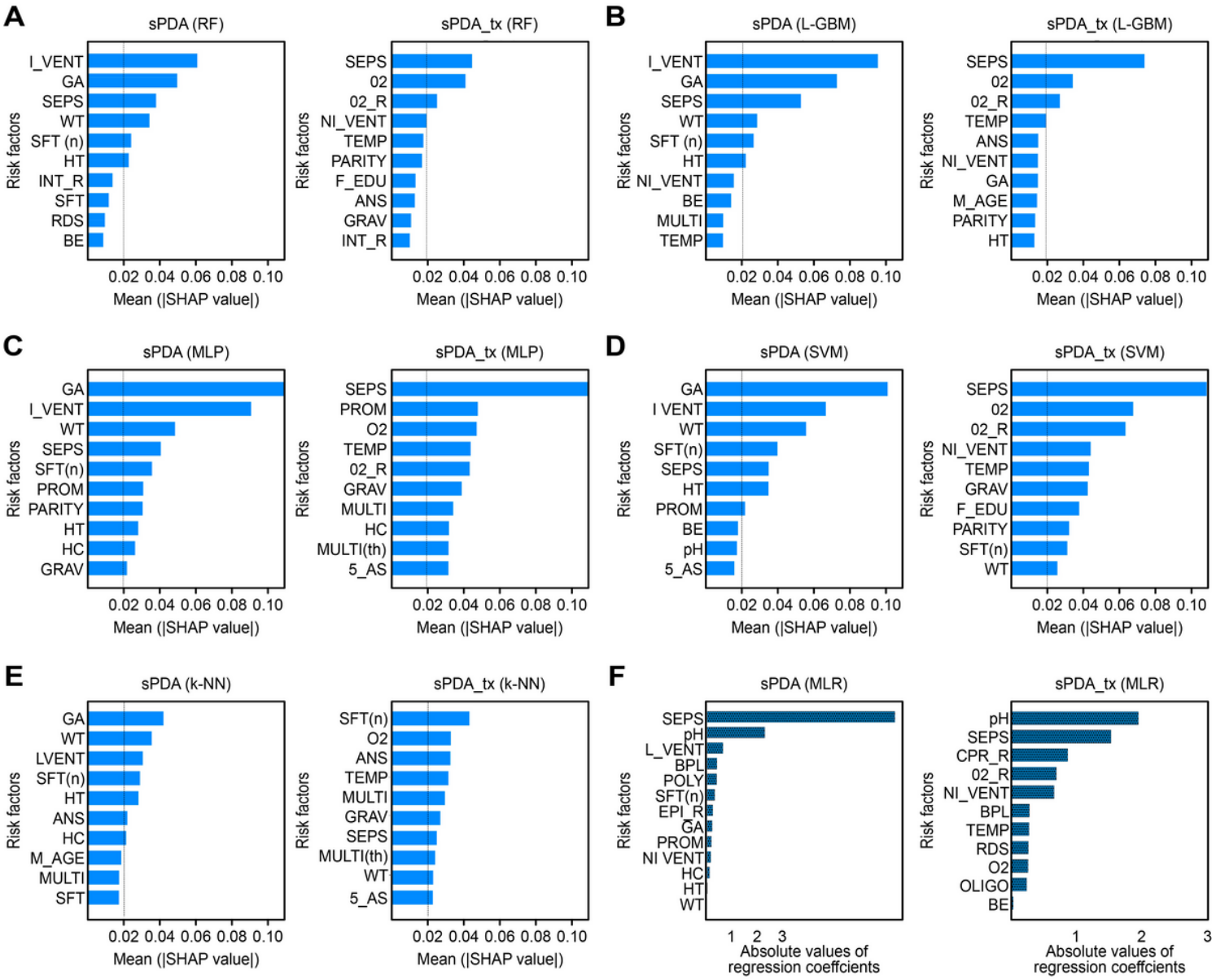

Figure 1

Top 10 factor contributions for sPDA and sPDA_tx prediction derived from each AA and MLR. a. The risk factors for sPDA and sPDA_tx prediction according to the RF. b. The risk factors for sPDA and sPDA_tx prediction according to the L-GBM. c. The risk factors for sPDA and PDA_tx prediction according to the MLP. d. The risk factors for sPDA and sPDA_tx prediction according to the SVM. e. The risk factors for sPDA and sPDA_tx prediction according to k-NN. The risk factors are listed in order of the average absolute SHAP values yielded by each algorithm in the artificial intelligence analysis and were selected based on a p-value of 0.05 during the testing procedure; the selected factors are sorted in descending order according to the absolute values of the corresponding regression coefficients in the MLR. Abbreviations: PDA, patent ductus arteriosus; sPDA, symptomatic PDA; sPDA_tx, sPDA with any treatment; RF, random forest; L-GBM, light gradient boosting machine; MLP, multilayer perceptron; SVM, support vector machine; k-NN, k-nearest neighbors; MLR, multiple logistic regression. The abbreviations for all factors are shown in Supplementary Table 1.

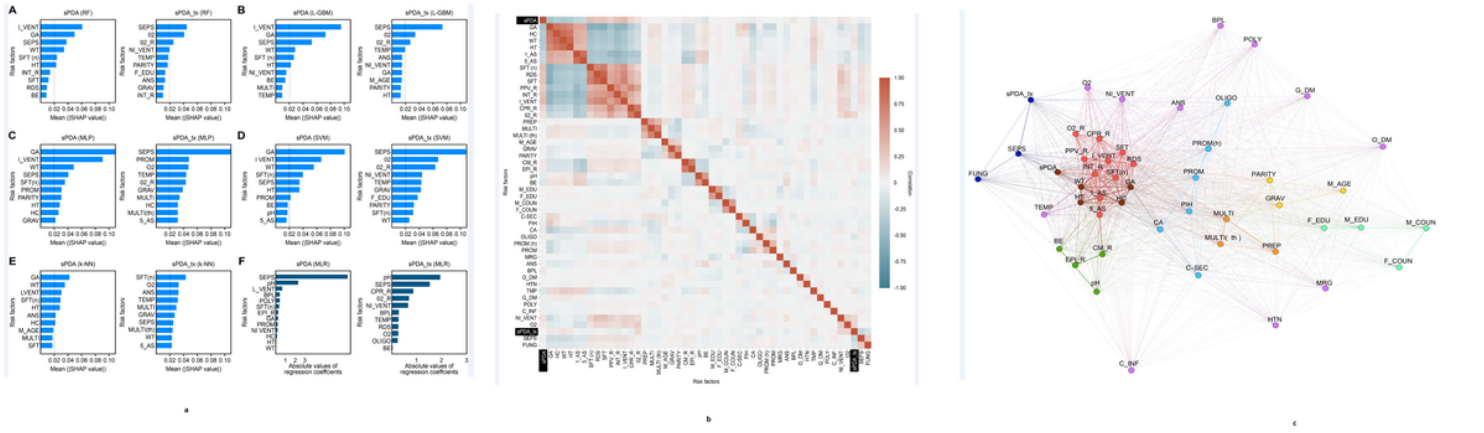

Figure 2

Relationships among the risk factors. a. A dendrogram visualizing hierarchical clustering based on the obtained correlation coefficients. The dendrogram's x-axis is composed of sPDA, sPDA_tx and all risk factors, and highly correlated factors are forced to be adjacent through hierarchical clustering. Each horizontal line indicates that the two associated subclusters are merged into one cluster, and the y-height indicates the distance between the two subclusters. We divided the factors into 9 clusters with a threshold of 1.15 and marked each cluster by color. b. A heatmap of the correlation matrix. The x-axis and y-axis of the heat map follow the arrangement of factors generated by hierarchical clustering, and the correlation coefficients are depicted in red or blue at the intersection of the factors. As the color bar on the right shows, red represents a positive correlation, and blue represents a negative correlation. A darker color indicates a higher correlation, while a lighter color indicates a lower correlation. c. A schematic diagram of the relationships among factors. The circles (nodes) represent the risk factors, and they are connected by the absolute value of the correlation coefficients (edges). In this network, the edges act as attraction forces, bringing highly correlated nodes closer together and pushing less-correlated nodes away from each other. The color of each cluster is the same as that in the dendrogram in Figure 2-a. Abbreviations: PDA, patent ductus arteriosus; sPDA, symptomatic PDA; sPDA_tx, sPDA with any treatment. The abbreviations for all factors are shown in Supplementary Table 1.

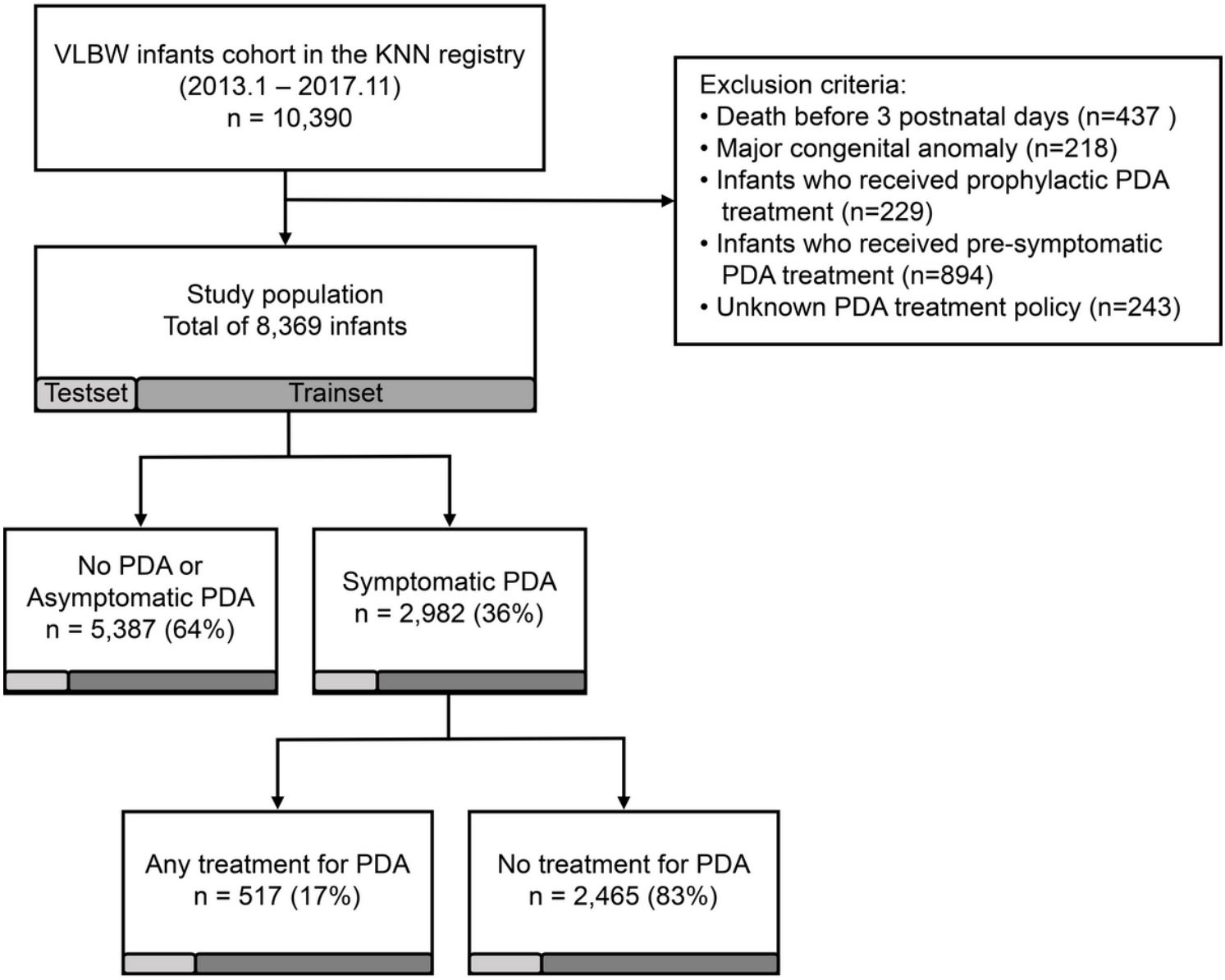

Figure 3

Study population identified with subsequent flowchart of the study. Abbreviations: VLBW infants, very low birth weight infants; KNN, Korean Neonatal Network; PDA, patent ductus arteriosus.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementmethods0831.docx
- Supplementtables0831.docx
- Supplementfigures0831.docx