Construction of Patient-level Prediction Model for In-hospital Mortality in Congenital Heart Disease Surgery: Regression and Machine Learning analysis

Xinwei Du
Shanghai Children's Medical Center, School of Medicine, Shanghai Jiaotong University
Hao Wang
Shanghai Children's Medical Center, School of Medicine, Shanghai Jiaotong University
Shunming Wang
Shanghai Children's Medical Center, School of Medicine, Shanghai Jiaotong University
Yi He
Shanghai Children's Medical Center, School of Medicine, Shanghai Jiaotong University
Jinghao Zheng
Shanghai Children's Medical Center, School of Medicine, Shanghai Jiaotong University
Haibo Zhang
Shanghai Children's Medical Center, School of Medicine, Shanghai Jiaotong University
Lifeng Yin
Synyi Medical Technology(Shanghai)
Yiwei Chen
Synyi Medical Technology (Shanghai)
Zhiwei Xu
Shanghai Children's Medical Center, School of Medicine, Shanghai Jiaotong University
Zhaohui Lu (Luzhaohui@scmc.com.cn)
Shanghai Children's Medical Center, School of Medicine, Shanghai Jiaotong University

---

Research

Keywords: Congenital heart disease, Procedure complexity score, In-hospital mortality, Predictive model, Machine learning

Posted Date: June 15th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-35146/v1

Abstract

Background: Prediction of in-hospital death is important for patient management as well as risk-adjusted evaluation of Congenital heart disease (CHD) surgery performance. Using a large database containing CHD surgery records of 12 years, we aim to establish patient-level in-hospital mortality prediction models.

Methods: Patients with congenital heart disease who underwent surgery at Shanghai Children's Medical Center from January 1, 2006, to December 31, 2017 were included in the study. Each procedure was assigned a complexity score based on Aristotle Score with modification. In-hospital mortalities for various surgery procedures were estimated. In-hospital death prediction models including a procedure complexity score and patient-level risk factors were constructed using logistic regression analysis and machine learning methods. The predictive values of the models were tested.

Results: Among 24,684 patients underwent CHD operations, there were 595 (2.4%) in-hospital deaths. The results showed that AUC of the prediction model based on logistic regression is 0.864 (95% CI: 0.833-0.895, P <0.001), the sensitivity is 0.831 and the specificity is 0.786. The AUC of the Gradient boosting model is 0.884 (95 %% CI: 0.858-0.909, P <0.001), the sensitivity and specificity were 0.838 and 0.785 respectively. The feature importance analysis found that the variable (average score) that had the greatest impact on the model's prediction performance was operation score (95.6), and other variables (average scores) were Age (days) (95.5), Ultrasound MV （54.6）, Ultrasound atrial level （54.5）, Palliative operation (45.8), Operation history (38.8), Ultrasound TV2 (32.1), Urgent operation (30.8), Ultrasound ventricular level (30.5), and Spo2 ≤ 90% (30.3).

Conclusions: Model constructed using machine learning method and logistic regression containing procedure complexity score and pre-operative patient-level factors had high accuracy in in-hospital mortality prediction. Operation score and age have the greatest impact on model prediction performance.

Background

Congenital heart disease (CHD) is the most common class of major congenital malformations. The prevalence of CHD at birth is estimated to be between 75 and 90 per 10 000 for live births and total pregnancies, with subgroup analyses showing rates of ≈ 1% for live births and 10% for aborted fetuses[1, 2]. It is also the leading cause of mortality from birth defects. Approximately one-quarter of the 40 000 children born with CHD annually in the United States require intervention in the first year of life [1, 3].

Since the first successful repair of a congenital heart defect using cardiopulmonary bypass in 1953, the accurate diagnosis and effective treatment of even the most complex congenital heart lesions have become standard practice. Surgery is offered for almost every heart defect, despite complexity [4]. Advances in the perioperative medical management of patients, particularly with intensive care, has also contributed to continuing improvement of CHD outcomes [5]. Early mortality for cardiac surgery in the neonatal period is ≈ 10% and beyond infancy is < 5%, with 90−95% of patients surviving with a good quality of life into the adult years [1]. However, there remains significant variability in outcomes according to procedures, patient characteristics, and other associated factors such as different hospitals.

In-hospital death is an important outcome of CHD. Precise in-hospital death prediction is of great importance in evaluating the quality of hospital surgical treatment, facilitating clinical decision on the performance of certain procedure, and improving hospital-patient relationship [6]. Since the complexity of procedure, as well as the risk factors at patient levels, are closely related with in-hospital death outcome of CHD surgery, case-mix adjustment is fundamental to any systematic attempt to measure outcomes, compare performance, and sustain a program of continual quality improvement. Two major methods of risk adjustment for CHD surgery, Risk Adjustment in Congenital Heart Surgery-1 (RACHS-1) [7], and the Aristotle Complexity Score [8] have been developed based on projections of risk or complexity that were predominantly subjectively derived. An objective, empirically based index to identify the statistically estimated risk of in-hospital mortality by procedure, the Society of Thoracic Surgery-European Association for Cardio-Thoracic Surgery (STAT) Congenital Heart Surgery Mortality score, have also been developed [9]. The feasibilities of these risk adjustment score have been tested in a few studies [10–12]. However, there still lacks study in Chinese population about the mortality risk prediction. In addition, these previous developed systems mainly aimed to evaluate the surgical outcomes among various hospitals. Also, due to the lack of information, these previous score systems concentrated more on surgical procedure categories but did not include sufficient risk factors at individual patient level. Therefore, the prediction may lack accuracy for individual patients.

Shanghai Children's Medical Center is a major tertiary hospital with advanced facility and teams in CHD treatment. A procedure score system has been adopted to stratify the in-hospital death risk of surgical procedures. The aim of current study is to establish and test a prediction model using the procedure score system together with risk factors at individual patient level.

Methods

Study design and population

Patients diagnosed with congenital heart disease and underwent surgery at Shanghai Children's Medical Center between January 1, 2006, and December 31, 2017 were included in the current study. For patients with multiple surgery records within 30 days, only the latest ones were used, the previous surgery records were considered as operation history. The exclusion criteria include general thoracic surgery not involving operation in heart, subjects with uncertain in-hospital survival records, and surgery procedure performed in < 3 patients. Demographic characteristics and pre-operation test data were extracted from a clinical database. In-hospital mortalities for various surgery procedures were estimated. The in-hospital death predictive value of a procedure risk stratification and patient-level risk factors were assessed with models constructed using logistic regression analysis and machine learning methods. The study was approved by the Ethical Committee of Shanghai Children's Medical Center. The inform consent was waived since the study only involved retrospective review of previous clinical data.

End Point

The study end point was in-hospital mortality, which was defined as death during the same hospitalization as surgery regardless of cause.

Data Source And Extraction

A “big data” database was constructed through the merging of multiple individual-level databases, included the hospital information system (HIS), laboratory information management system (LIS), clinical data repository (CDR), intensive care unit (ICU) database, and surgery record database of cardiac surgical department in Shanghai Children’s Medical Center. These datasets were linked using unique encoded personal identifiers. To merge and standardize the data from multiple sources, semantic transformation process was handled with ETL approach. Data extracted from the original sources were provided as input to the semantic transformation rules which are created by domain experts to express semantics of the data transformation. Data were transformed according to a standardized common data module (CDM).

Baseline characteristics including demographic information, diagnosis, pre-operation results of echocardiography and laboratory tests were extracted from merged database. Variables with missing valued in more than 30% cases were excluded.

Statistical Analysis And Model Construction

Continuous variables were described using Median (Range), since all of them are not normally distributed. Categorical variables were described using frequency (%). To evaluate the distributive balance between training set and validation set, Mann-Whitney U test, chi-squared test and Fisher’s exact test were applied for between groups comparison, as appropriate.

Mortality of each procedure was calculated. In addition, estimations of procedure-specific mortality rates were calculated by using a Bayesian random effects model that adjusted each procedure’s mortality rate based on the size of the denominator. Unlike conventional methods, random effects models use data from all the procedures in the database when estimating the probability of mortality for any single procedure. This “borrowing of information” across procedures produces estimates with good statistical properties, including smaller standard errors than conventional estimates. The model-based estimate is a weighted average of a procedure’s actual observed mortality rate and the overall average mortality rate for all procedures in the database. The model weights an individual procedure’s own data more heavily when the denominator is large enough to be reliable and weights the overall average mortality rate more heavily when the denominator is too small to support a reliable mortality estimate.

A mortality risk stratification was performed by categorizing procedures into groups according to mortality. Performance of different categorizations consisting of 2 to 20 categories were evaluated based on the internal homogeneity of the categories and the discrimination of the categories as predictors of mortality.

To construct in-hospital death prediction model at patient level, we firstly conducted univariate analysis on training set with aim to selection of potential risk factors. Then still in training set, we built both multiple logistic regression and gradient boosting decision tree (GBDT) model based on the potential risk factors found in univariate analysis. Specifically, GBDT is an ensemble algorithm for classification and regression with multiple tree models under gradient boosting, and we developed the machine learning model with the XGBoost toolkit, an implementation of the GBDT algorithm with strong flexibility. A grid search was used for the setting of the XGBoost model hyper-parameters[13]. Finally, we applied the constructed models (i.e. logistic regression and XGBoost model) in validation set, we used ROC curve analysis to evaluate the predictive power of the models. Specifically, AUC together with 95% confidence and P value, sensitivity and specificity were calculated. In addition, importance of top 10 risk factors in prediction of XGBoost model were also produced.

All statistical tests were two-sided tests, $P < 0.05$ was considered statistically significant. All statistical analysis and the logistic regression were conducted with SAS (version xx?). The XGBoost model was implemented with Python (version 3.6) library XGBoost (version 0.9)

Results

Characteristics of study population

As shown in Table 1, a total of 24,684 patients underwent CHD operations were included, among those were 595 (2.4%) in-hospital deaths. Demographic and pre-operation characteristics of patients with death or survival outcomes were showed in Supplemental Table 1. The most commonly performed procedures included VSD repair, tetralogy repair, ASD secundum repair etc. The in-hospital mortality for each procedure varies between 0 to 77.78%, with 24 procedures with 0 death record. When constructing a death risk prediction model, 75% of patients were included in the training set and 25% of patients were included in the validation set. There were no significant differences between most of the demographic and pre-operation characteristics of the training set and validation set, except for white blood cell count ($P = 0.02$), basophil count ($P = 0.03$), and MPV ($P = 0.004$) between training set and validation set (supplemental table 2).

Table 1
In-hospital mortality of individual procedure performed

Procedure	Cases (n)	Death (n)	Mortality (unadjusted)
Decannulation	9	7	77.78%
ECMO	18	13	72.22%
Fontan Takedown	9	5	55.56%
Rt. or lt. heart assist	8	3	37.50%
Atrial septectomy	12	4	33.33%
Coronary artery repair	6	2	33.33%
Double switch (Senning + ASO)	17	5	29.41%
DKS connection	7	2	28.57%
Truncus repair	33	7	21.21%
Central shunt - with graft	38	8	21.05%
Arterial switch repair	451	83	18.40%
Conduit RV - PA	71	13	18.31%
PA unifocalization	61	11	18.03%
R.E.V. RV - PA connection	7	1	14.29%
TAPVD repair - infracardiac	43	6	13.95%
Cavopulmonary shunt - left	83	11	13.25%
Interrupted aortic arch repair	76	10	13.16%
ALC-PA repair	55	7	12.73%
Interruption of bronchial collaterals	8	1	12.50%
Other procedures	49	6	12.24%
Rastelli operation	50	6	12%
PA banding	101	12	11.88%
Ao translocation operation	17	2	11.76%
Pacemaker - primary implant	17	2	11.76%
Mitral replacement (mech.)	95	10	10.53%
Pulmonary atresia/IVS repair	19	2	10.53%

Procedure	Cases (n)	Death (n)	Mortality (unadjusted)
Cavopulmonary shunt - right	359	35	9.75%
Aortic arch repair	21	2	9.52%
Takedown previous shunt	11	1	9.09%
Aortic valve replacement (mech.)	44	3	6.82%
Pulmonary atresia/VSD repair	288	17	5.90%
Cavopulmonary shunt - bilateral	102	6	5.88%
Delayed sternal closure	891	51	5.72%
Fontan operation - II stage	537	30	5.59%
AVSD repair - two patches	269	15	5.58%
Pulmonary valvotomy (off pump)	18	1	5.56%
TAPVD repair - supracardiac	238	13	5.46%
Senning procedure	19	1	5.26%
Pulm. Infundibulum resection, patch across annulus	292	15	5.14%
Pulm. Infundibulum resection & patch	78	4	5.13%
Pericardial drainage	160	8	5%
Pulm. vein stenosis repair	40	2	5%
Aortic arch repair	48	2	4.17%
Fontan operation - I stage	110	4	3.64%
Asc. Aorta Patch aortoplasty	183	6	3.28%
A-P window Repair	31	1	3.23%
DORV repair – IVR	375	12	3.20%
Pulmonary arterioplasty	251	8	3.19%
TAPVD repair - mixed type	34	1	2.94%
Coarct. repair - resection & E to S	376	9	2.39%
Pulm. Infundibulum incision & resection	46	1	2.17%
Excision of cardiac tumor	96	2	2.08%
Tetralogy repair	2390	45	1.88%

Procedure	Cases (n)	Death (n)	Mortality (unadjusted)
Coarct. repair - patch aortoplasty	253	4	1.58%
VSD multiple repair	127	2	1.57%
PDA closure (off pump)	386	6	1.55%
Mitral valvuloplasty	461	7	1.52%
Pulmonary valvotomy	200	3	1.50%
Subaortic myectomy	80	1	1.25%
TAPVD repair - intracardiac	184	2	1.09%
VSD muscular repair	105	1	0.95%
Ebstein anomaly repair	106	1	0.94%
Aortic valvuloplasty	109	1	0.92%
AVSD repair - single patch	116	1	0.86%
AVSD repair - depression	390	3	0.77%
VSD Hybrid repair (off pump)	171	1	0.58%
Tricuspid valvuloplasty	540	3	0.56%
VSD membranous repair	7907	22	0.28%
ASD secundum repair	550	1	0.18%
ASD secundum repair (patch)	1338	1	0.07%
VSD subarterial repair	1641	1	0.06%
ASD Hybrid repair (off pump)	87	0	0%
ASD common atrium repair	9	0	0%
ASD repair - minimal invasive & CPB	29	0	0%
ASD sinus venosus repair	38	0	0%
Aortic valvotomy	56	0	0%
Coronary art. fistula repair	78	0	0%
Cortraitriatum repair	70	0	0%
Exicion of intracardiac vegetation	11	0	0%
Hemitruncus repair	21	0	0%

Procedure	Cases (n)	Death (n)	Mortality (unadjusted)
PAPVD isolated repair	115	0	0%
PDA closure (CPB)	32	0	0%
Pericardectomy	20	0	0%
Pulm. Infundibulum resection (indirect)	178	0	0%
Pulmonary art. sling repair	20	0	0%
Pulmonary art. stent	13	0	0%
Subaortic fibromyectomy	155	0	0%
Subaortic septal patch (Konno)	14	0	0%
Supravalve mitral ring resection	28	0	0%
Systemic vein repair	14	0	0%
Tricuspid replacement (mech.)	13	0	0%
VSD canal type repair	154	0	0%
VSD hybrid repair (CPB)	21	0	0%
VSD repair - minimal invasive & CPB	45	0	0%
Vascular ring repair	132	0	0%

Table 1
In-hospital mortality of individual procedure performed (sorted by mortality rate)

Procedure	Cases (n)	Death (n)	Mortality (unadjusted)
Decannulation	9	7	77.78%
ECMO	18	13	72.22%
Fontan Takedown	9	5	55.56%
Rt. or lt. heart assist	8	3	37.50%
Atrial septectomy	12	4	33.33%
Coronary artery repair	6	2	33.33%
Double switch (Senning + ASO)	17	5	29.41%
DKS connection	7	2	28.57%
Truncus repair	33	7	21.21%
Central shunt - with graft	38	8	21.05%
Arterial switch repair	451	83	18.40%
Conduit RV - PA	71	13	18.31%
PA unifocalization	61	11	18.03%
R.E.V. RV - PA connection	7	1	14.29%
TAPVD repair - infracardiac	43	6	13.95%
Cavopulmonary shunt - left	83	11	13.25%
Interrupted aortic arch repair	76	10	13.16%
ALC-PA repair	55	7	12.73%
Interruption of bronchial collaterals	8	1	12.50%
Other procedures	49	6	12.24%
Rastelli operation	50	6	12%
PA banding	101	12	11.88%
Ao translocation operation	17	2	11.76%
Pacemaker - primary implant	17	2	11.76%
Mitral replacement (mech.)	95	10	10.53%
Pulmonary atresia/IVS repair	19	2	10.53%

Procedure	Cases (n)	Death (n)	Mortality (unadjusted)
Cavopulmonary shunt - right	359	35	9.75%
Aortic arch repair	21	2	9.52%
Takedown previous shunt	11	1	9.09%
Aortic valve replacement (mech.)	44	3	6.82%
Pulmonary atresia/VSD repair	288	17	5.90%
Cavopulmonary shunt - bilateral	102	6	5.88%
Delayed sternal closure	891	51	5.72%
Fontan operation - II stage	537	30	5.59%
AVSD repair - two patches	269	15	5.58%
Pulmonary valvotomy(off pump)	18	1	5.56%
TAPVD repair - supracardiac	238	13	5.46%
Senning procedure	19	1	5.26%
Pulm. Infundibulum resection, patch across annulus	292	15	5.14%
Pulm. Infundibulum resection & patch	78	4	5.13%
Pericardial drainage	160	8	5%
Pulm. vein stenosis repair	40	2	5%
Aortic arch repair	48	2	4.17%
Fontan operation - I stage	110	4	3.64%
Asc. Aorta Patch aortoplasty	183	6	3.28%
A-P window Repair	31	1	3.23%
DORV repair – IVR	375	12	3.20%
Pulmonary arterioplasty	251	8	3.19%
TAPVD repair - mixed type	34	1	2.94%
Coarct. repair - resection & E to S	376	9	2.39%
Pulm. Infundibulum incision & resection	46	1	2.17%
Excision of cardiac tumor	96	2	2.08%
Tetralogy repair	2390	45	1.88%

Procedure	Cases (n)	Death (n)	Mortality (unadjusted)
Coarct. repair - patch aortoplasty	253	4	1.58%
VSD multiple repair	127	2	1.57%
PDA closure (off pump)	386	6	1.55%
Mitral valvuloplasty	461	7	1.52%
Pulmonary valvotomy	200	3	1.50%
Subaortic myectomy	80	1	1.25%
TAPVD repair - intracardiac	184	2	1.09%
VSD muscular repair	105	1	0.95%
Ebstein anomaly repair	106	1	0.94%
Aortic valvuloplasty	109	1	0.92%
AVSD repair - single patch	116	1	0.86%
AVSD repair - depression	390	3	0.77%
VSD Hybrid repair (off pump)	171	1	0.58%
Tricuspid valvuloplasty	540	3	0.56%
VSD membranous repair	7907	22	0.28%
ASD secundum repair	550	1	0.18%
ASD secundum repair (patch)	1338	1	0.07%
VSD subarterial repair	1641	1	0.06%
ASD Hybrid repair (off pump)	87	0	0%
ASD common atrium repair	9	0	0%
ASD repair - minimal invasive & CPB	29	0	0%
ASD sinus venosus repair	38	0	0%
Aortic valvotomy	56	0	0%
Coronary art. fistula repair	78	0	0%
Cortraitriatum repair	70	0	0%
Exicion of intracardiac vegetation	11	0	0%
Hemitruncus repair	21	0	0%

Procedure	Cases (n)	Death (n)	Mortality (unadjusted)
PAPVD isolated repair	115	0	0%
PDA closure (CPB)	32	0	0%
Pericardectomy	20	0	0%
Pulm. Infundibulum resection (indirect)	178	0	0%
Pulmonary art. sling repair	20	0	0%
Pulmonary art. stent	13	0	0%
Subaortic fibromyectomy	155	0	0%
Subaortic septal patch (Konno)	14	0	0%
Supravalve mitral ring resection	28	0	0%
Systemic vein repair	14	0	0%
Tricuspid replacement (mech.)	13	0	0%
VSD canal type repair	154	0	0%
VSD hybrid repair (CPB)	21	0	0%
VSD repair - minimal invasive & CPB	45	0	0%
Vascular ring repair	132	0	0%

Stratification Of Surgical Procedure

Previous analysis has clearly shown that the complexity of procedure is among the most important factors in predicting in-hospital death. To further categorize the mortality risk of various procedures, performance of different categorizations consisting of 2 to 20 categories were evaluated based on the internal homogeneity of the categories and the discrimination of the categories as predictors of mortality. Procedures were sorted by increasing estimated risk and partitioned into 6 relatively homogeneous categories (Table 2). Among all the 95 surgical methods included in the analysis, there were 3 types of surgery with a risk stratification of level 6, 10 with a stratification of level 5, 11 with level 4, 19 with level 3, and 2 There are 38 types and 14 types for level 1. In our study, surgery risk stratified by level 6 includes Decannulation, ECMO, and Fontan Takedown, with corresponding mortality rates of 71.55%, 64.99%, and 52.15%, which are higher than other surgical risk levels (Table2).

Table 2
Risk stratification of procedures based on actual mortality analysis

Procedure	Mortality (model based)	Risk stratification
ASD Hybrid repair (off pump)	0.45%	1
ASD secundum repair	0.12%	1
ASD secundum repair (patch)	0.14%	1
Aortic valvuloplasty	0.36%	1
Coronary art. fistula repair	0.44%	1
Cortraitriatum repair	0.45%	1
PAPVD isolated repair	0.35%	1
Pulm. Infundibulum resection (indirect)	0.26%	1
Subaortic fibromyectomy	0.28%	1
Tricuspid valvuloplasty	0.30%	1
VSD canal type repair	0.30%	1
VSD membranous repair	0.22%	1
VSD subarterial repair	0.12%	1
Vascular ring repair	0.32%	1
ASD common atrium repair	1.17%	2
ASD repair - minimal invasive & CPB	0.84%	2
ASD sinus venosus repair	0.76%	2
AVSD repair - depression	0.98%	2
AVSD repair - single patch	1.09%	2
Aortic valvotomy	0.51%	2
Cavopulmonary shunt - bilateral	2.39%	2
Coarct. repair - resection & E to S	2.04%	2
Coarct. repair - patch aortoplasty	1.99%	2
DORV repair – IVR	2.93%	2
Ebstein anomaly repair	1.10%	2
Excision of cardiac tumor	2.72%	2
Exicion of intracardiac vegetation	1.23%	2

Procedure	Mortality (model based)	Risk stratification
Fontan operation - I stage	2.10%	2
Hemitruncus repair	0.84%	2
Interruption of bronchial collaterals	1.30%	2
Mitral valvuloplasty	1.43%	2
PDA closure (CPB)	0.71%	2
PDA closure (off pump)	1.98%	2
Pericardectomy	0.86%	2
Pulm. Infundibulum incision & resection	2.33%	2
Pulmonary art. sling repair	0.84%	2
Pulmonary art. stent	1.06%	2
Pulmonary arterioplasty	2.65%	2
Pulmonary valvotomy	1.75%	2
Subaortic myectomy	1.47%	2
Subaortic septal patch (Konno)	0.98%	2
Supravalve mitral ring resection	0.82%	2
Systemic vein repair	0.98%	2
TAPVD repair - intracardiac	1.36%	2
Takedown previous shunt	1.06%	2
Tetralogy repair	1.93%	2
Tricuspid replacement (mech.)	1.02%	2
VSD Hybrid repair (off pump)	0.77%	2
VSD hybrid repair (CPB)	0.86%	2
VSD multiple repair	0.97%	2
VSD muscular repair	1.29%	2
VSD repair - minimal invasive & CPB	0.56%	2
A-P window Repair	3.06%	3
AVSD repair - two patches	5.73%	3
Ao translocation operation	4.61%	3

Procedure	Mortality (model based)	Risk stratification
Aortic arch repair	4.80%	3
Aortic valve replacement (mech.)	5.21%	3
Asc. Aorta Patch aortoplasty	4.11%	3
Delayed sternal closure	5.42%	3
Fontan operation - II stage	5.57%	3
Pacemaker - primary implant	4.61%	3
Pericardial drainage	5.30%	3
Pulm. Infundibulum resection & patch	4.39%	3
Pulm. Infundibulum resection, patch across annulus	4.29%	3
Pulm. vein stenosis repair	5.38%	3
Pulmonary atresia/IVS repair	4.61%	3
Pulmonary atresia/VSD repair	4.99%	3
Pulmonary valvotomy (off pump)	5.15%	3
Senning procedure	4.61%	3
TAPVD repair - mixed type	3.06%	3
TAPVD repair - supracardiac	5.59%	3
ALC-PA repair	12.07%	4
Aortic arch repair	9.04%	4
Cavopulmonary shunt - left	9.93%	4
Cavopulmonary shunt - right	8.32%	4
DKS connection	8.64%	4
Mitral replacement (mech.)	10.09%	4
PA banding	10.29%	4
R.E.V. RV - PA connection	8.64%	4
Rastelli operation	12.72%	4
TAPVD repair - infracardiac	12.72%	4
Other procedures	7.92%	4
Arterial switch repair	19.65%	5

Procedure	Mortality (model based)	Risk stratification
Atrial septectomy	31.67%	5
Central shunt - with graft	24.14%	5
Conduit RV - PA	19.79%	5
Coronary artery repair	21.49%	5
Double switch (Senning + ASO)	26.91%	5
Interrupted aortic arch repair	14.03%	5
PA unifocalization	15.49%	5
Rt. or lt. heart assist	21.49%	5
Truncus repair	15.33%	5
Decannulation	71.55%	6
ECMO	64.99%	6
Fontan Takedown	52.15%	6

Univariate Analysis For Risk Factors Of In-hospital Mortality

We performed a univariate analysis of the demographic and preoperative characteristics of the dead and surviving patients in the training and validation sets, and then statistically significant variables were included into the model. From the results of the single factor analysis, the variables included in the model include age, MV, TV2, pre-operation hct, pre-operation Spo2, pre-operation special treatment, CPB, urgent operation, palliative surgery, previous surgery, aortic level shunt, atrial level shunt, ventricular level shunt, PI, AO, LA, LVDd, operation risk stratification, white blood cell count, monocyte count, neutrophil count, basophil count, red blood cell count, PCV (Hct), MCV, MCHC (Mass), platelet count, MPV, APTT, PT, total protein, serum albumin, A/G, ALT, AST, ALP, total bilirubin, direct bilirubin, creatinine, uric acid, eosinophils/100 white blood cells, monocytes/100 white blood cells, and basophils/100 white blood cells.

Construction of in-hospital death prediction model based on patient level

Patient level risk factors including baseline characteristics, pre-operation test results of echocardiography and laboratory tests, as well as the procedure score were compared between populations with death and survival outcomes. The results identified significant differences in numerous factors between the two groups (supplemental Table 2).

In concerning that the numerous pre-operation data incorporated into regression analysis may interfere with the effectiveness of the prediction model, a different approach with machine learning model construction, which is considered to be more effective in “big data” analysis, was performed for in-hospital death prediction. The preoperative indicators included basic information, medical history, diagnostic categories, and pre-operation test results.

The data extracted for the study samples contained 124 potential risk factor features. In this experiment, we conducted a one-by-one examination of each feature and removed 51 features with more than 30% data missing. A normal range (at least 95% of the samples were included) and a proper transform function was assigned to each continuous feature and values beyond the reasonable range was set as boundary values to increase the effectiveness of the experiment. Binary data was converting to value 0 and 1. Ranking data were assigned into ascending numbers. Categorical data were encoded into individual features containing binary values. Missing data were filled with the median (mode) value for each continuous (discrete) feature.

The data processing resulted in 158 risk factor features in the final model analysis.

Receiver operating characteristic curves for the risk prediction value of model based on logistic regression and gradient boosting model were displayed in Table 3 and Fig. 1. The AUC of Logistic regression prediction model was 0.864 (95% CI: 0.833, 0.895), $P < 0.001$. The sensitivity is 0.831 and the specificity is 0.786. The AUC of the Gradient boosting model is 0.884 (95%% CI: 0.858–0.909, $P < 0.001$), the sensitivity and specificity were 0.838 and 0.785 respectively (Table 3).

Table 3
Prediction value of model based on logistic regression and gradient boosting model

Model	AUC	95% CI	P value	Sensitivity	Specificity
Logistic regression	0.864	(0.833, 0.895)	< 0.001	0.831	0.786
Gradient boosting model	0.884	(0.858, 0.909)	< 0.001	0.838	0.785

Importance of top 10 risk factors in prediction of gradient boosting model

The importance of top 10 risk factors in prediction of gradient boosting model were analyzed (Table 4). The results showed that the average score of operation score, Age (days), Ultrasound MV, Ultrasound atrial level, Palliative operation, Operation history, Ultrasound TV2, Urgent operation, Ultrasound ventricular level, and Spo2 ≤ 90% were 95.6, 95.5, 54.6, 54.5, 45.8, 38.8, 32.1, 30.8, 30.5, and 30.3 respectively.

Table 4
Importance of top 10 risk factors in prediction of gradient boosting model

Feature	Average score
Operation score	95.6
Age (days)	95.5
Ultrasound MV	54.6
Ultrasound atrial level	54.5
Palliative operation	45.8
Operation history	38.8
Ultrasound TV2	32.1
Urgent operation	30.8
Ultrasound ventricular level	30.5
Spo2 $\leq$ 90%	30.3

A higher score of a feature stands for a more significant role in the machine learning model for outcome classification. The risk factor features that have the greatest impact on the model are the procedure score, age, Ultrasound MV, and Ultrasound atrial level.

Discussion

Assessing the surgical risk of cardiac surgery is critical to understanding the difference in surgical outcomes, including in-hospital mortality and improving patient outcomes. In-hospital mortality risk prediction has important clinical significance for assessing the quality of cardiac surgery and patient postoperative management in different institutions.

In our study, in-hospital mortality risk prediction models were constructed based on combination of surgical risk stratification and patient pre-operation variables, and the predictive power of the models for postoperative mortality risk was evaluated. The results showed that in-hospital mortality of CHD surgery in our hospital was 2.4%, which was consistent with previously reported results in Western countries [14–16] and much lower than studies in developing country [10]. This mortality was also similar as the previous overall mortality of CHD patients in China [17]. According to the stratification of RACHS-1, the mortality of CHD procedures with different risk varied from 0.26–62% [12, 18], which was also consistent with the mortality distribution observed in the current study.

Our research found that operation score has the greatest impact on the predictive performance of the death risk prediction model. In our study, Decannulation, ECMO, and Fontan Takedown with the highest surgical risk scores (6 points) also had the highest mortality rates, 77.78%, 72.22%, and 55.56%, respectively. Several surgery risk stratification systems have been developed previously, such as Aristotle score and The Risk Adjustment for Congenital Heart Surgery (RACHS- 1) system [19].[20]. The predictive power of Aristotle comprehensive complexity was better than Aristotle basic complexity and risk adjustment in congenital heart surgery-1 (RACHS-1) prediction models for in hospital mortality after surgery for congenital heart disease[21]. Hörer J analyzed the predictive power of the ACHS score for mortality risk in adults after congenital heart surgery, and they found that the ACHS score had similar, and good predictive power for surgical outcomes in 2 European centers[22]. It has been reported that the mortality of CHD procedures with different risk stratification of RACHS-1 varied, with higher mortality in higher risk level groups[12, 18]. Another study reported mortality of 8.76% (48/548) for categories 1 and 2 pooled and 26.12% (70/268) to the categories 3 and 4[23]. Based on STS-EACTS stratification method, the hospital mortality was 33% for more complex procedures (categories 3 and 4) against 7.98% for the less complex (categories 1 and 2) [12].

Due to enormous variability in CHD procedures as well as patient-level factors, it is difficult to establish a predictive model for mortality. All these models mentioned above were constructed as a risk-adjustment tool for of performance evaluation of CHD surgeries. In China, there still lack uniform standards for administration and evaluation of CHD surgery.

The Aristotle Score system was adopted with modification according to our experiences. The individual procedure scoring system instead of risk stratification level was chosen in consideration of the request of an individual instead of group level predictive model. As tools to compare surgery performance across groups, the complexity score/stratification had certain prediction value however not very accurate. It has been reported that when used for prediction of in-hospital death, the ABC score, RACHS-1 and STS-EACTs were usually with an overall C-index from 0.6–0.8, varies in different populations [10, 21, 24–27].

In addition to surgical risk stratification, patient-level factors also have a significant impact on the predictive power of predictive models for postoperative mortality risk. A clinical study found that, the mortality risk prediction model constructed based on hemodynamic support, cardiac diagnosis, renal dysfunction and total bilirubin, had a good power for prediction of the in-hospital mortality after children undergoing heart transplantation, and might ay provide support for clinical decision-making of cardiac transplantation and the use of mechanical ventilation[28]. In our study, large numbers of pre-operative patient-level factors were selected to construct the model. Both logistic regression and machine learning methods have incorporated age, pre-operative echocardiography characteristics and lab examination results into the final predictive model. The predictive values (AUC) of the final models were 0.864 and 0.884 separately, which was higher than previous reported for ABC score, RACHS-1 and STS-EACTs, even after incorporation of patient-level factors[21, 26, 29].

Artificial intelligence and machine learning model has been rapidly incorporated in “big data analysis” [30–32]. In a retrospective study, based on the data from a single tertiary center including 10019 patients, a machine learning algorithm was used to build a model to predict the prognosis of adult patients with ACHD and evaluate its guiding role in clinical treatment. The results showed accuracy of deep learning model based on categorized diagnosis, disease complexity, and NYHA class were 91.1%, 97.0%, and 90.6%, respectively in the test sample[33]. A multicenter retrospective cohort study found echocardiography-based deep learning model are more accurate in predicting the risk of in-hospital mortality among patients with heart disease than other models[34]. However, in China, there is currently no relevant research on the risk prediction model based on machine learning to predict in-hospital mortality risk after surgery in children with congenital heart disease. In the current study, both traditional statistical analysis with logistic regression and machine learning methods have been tried in the model construction. Considering that the machine learning model was constructed based on a step-by-step validation principle, while there was no validation cohort in this study for the regression model, we speculate that the machine learning model construction may be with higher practical value and this method may further approve the future model construction / modification.

The current study was with certain limitations: As all studies in CHD surgery risk evaluation, due to the great variations in CHD surgery procedures and low mortality, the sample size of individual procedure was small for most procedures. The statistical methods used to adjust may not be sufficient to overcome the influence. In addition, the data used in the current study was from a single center, which might not be representative for the general practice in China. The experiences with specific procedure may brought bias into the analysis. However, this was the first systematic analysis of CHD procedure risk in China with a large sample size. The results should, at least in part, reflect the current status in CHD surgery practice in China. Also, due to the limited data source, especially the low mortality, we did not include a validation cohort in the current study, therefore the real value of the prediction models needs to be further validated in future studies. Finally, as previous studies, the in-hospital mortality recorded may not be representative for all operation-related death, which should incorporate data in a 30 days period after discharge. This needs to be addressed in a more complete data source.

Conclusion

In conclusion, the current study has demonstrated that through combination of procedure complexity score with pre-operative patient-level factors, model constructed using machine learning method and logistic regression had high but similar accuracy in in-hospital mortality prediction. Operation score and age have the greatest impact on model prediction performance. The predictive model may be applied for surgery performance evaluation as well as pre-operative risk prediction in clinical practice.

Abbreviations

CHD
Congenital heart disease
RACHS-1
Risk Adjustment in Congenital Heart Surgery-1
STS
Society of Thoracic Surgeons

EACTS
European Congenital Heart Surgery Databases of Association for Cardiothoracic Surgery
CDM
common data module
HIS
Hospital Information System
LIS
Laboratory Information Management System
CDR
Clinical Data Repository
ICU
intensive care unit
ROC
receiver operating characteristic

Declarations

Ethics approval and consent to participate

The study was approved by the Ethical Committee of Shanghai Children's Medical Center. The inform consent was waived since the study only involved retrospective review of previous clinical data.

Consent for publication

Not applicable

Availability of data and materials

Due to hospital policy restrictions, the datasets used and/or analyzed during the current study are not available.

Competing interests

All authors declared that there were no financial and personal relationships with other people or organizations that could inappropriately influence (bias) their work.

Funding

None.

Authors' contributions

Xinwei Du: responsible for data collection and data review, data cleaning

Hao Wang: responsible for research implementation, data collection and analysis

Shunmin Wang: responsible for supervision of project execution, data quality control

Yi He: accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved

Jinghao Zheng: data review, and revising the article

Haibo Zhang: responsible for data collection, data quality control, revising the article

Jinfen Liu: responsible for data collection, revising the article

Lifeng Yin: responsible for the statistics and analysis of research data

Yiwei Chen: responsible for retrieval and screening literature, writing research papers

Zhiwei Xu: data review, final approval of the version to be published

Zhaohui Lu: responsible for the determination of the research direction, the design of the research program, and the summary of the research questions

Acknowledgments

Thanks to all the staff of Shanghai Children's Medical Center who have contributed to the study. Thanks to Shanghai Synyi Medical Technology Co., Ltd for providing a data analysis and statistical platform.

References

1. 1. Triedman JK, Newburger JW: **Trends in Congenital Heart Disease: The Next Decade.** *Circulation* 2016, **133**:2716–2733.
2. 2. Kahr PC, Diller GP: **Almanac 2014: congenital heart disease.** *Heart* 2015, **101**:65–71.
3. 3. van der Bom T, Zomer AC, Zwinderman AH, Meijboom FJ, Bouma BJ, Mulder BJ: **The changing epidemiology of congenital heart disease.** *Nat Rev Cardiol* 2011, **8**:50–60.
4. 4. Bouma BJ, Mulder BJ: **Changing Landscape of Congenital Heart Disease.** *Circ Res* 2017, **120**:908–922.
5. 5. Holst KA, Said SM, Nelson TJ, Cannon BC, Dearani JA: **Current Interventional and Surgical Management of Congenital Heart Disease: Specific Focus on Valvular Disease and Cardiac Arrhythmias.** *Circ Res* 2017, **120**:1027–1044.
6. 6. Berger JT, Holubkov R, Reeder R, Wessel DL, Meert K, Berg RA, Bell MJ, Tamburro R, Dean JM, Pollack MM: **Morbidity and mortality prediction in pediatric heart surgery: Physiological profiles and surgical complexity.** *J Thorac Cardiovasc Surg* 2017, **154**:620–628.e626.

7. 7. Jenkins KJ, Gauvreau K: **Center-specific differences in mortality: preliminary analyses using the Risk Adjustment in Congenital Heart Surgery (RACHS-1) method.** *J Thorac Cardiovasc Surg* 2002, **124**:97–104.
8. 8. Lacour-Gayet F, Clarke D, Jacobs J, Gaynor W, Hamilton L, Jacobs M, Maruszewski B, Pozzi M, Spray T, Tchervenkov C, Mavroudis C: **The Aristotle score for congenital heart surgery.** *Semin Thorac Cardiovasc Surg Pediatr Card Surg Annu* 2004, **7**:185–191.
9. 9. O'Brien SM, Clarke DR, Jacobs JP, Jacobs ML, Lacour-Gayet FG, Pizarro C, Welke KF, Maruszewski B, Tobota Z, Miller WJ, et al: **An empirically based tool for analyzing mortality associated with congenital heart surgery.** *J Thorac Cardiovasc Surg* 2009, **138**:1139–1153.
10. 10. Cavalcante CT, de Souza NMG, Pinto VCJ, Branco KM, Pompeu RG, Teles AC, Cavalcante RC, de Andrade GV: **Analysis of Surgical Mortality for Congenital Heart Defects Using RACHS-1 Risk Score in a Brazilian Single Center.** *Braz J Cardiovasc Surg* 2016, **31**:219–225.
11. 11. Miyata H, Murakami A, Tomotaki A, Takaoka T, Konuma T, Matsumura G, Sano S, Takamoto S: **Predictors of 90-day mortality after congenital heart surgery: the first report of risk models from a Japanese database.** *J Thorac Cardiovasc Surg* 2014, **148**:2201–2206.
12. 12. Cavalcanti PE, Sa MP, Santos CA, Esmeraldo IM, Chaves ML, Lins RF, Lima Rde C: **Stratification of complexity in congenital heart surgery: comparative study of the Risk Adjustment for Congenital Heart Surgery (RACHS-1) method, Aristotle basic score and Society of Thoracic Surgeons-European Association for Cardio- Thoracic Surgery (STS-EACTS) mortality score.** *Rev Bras Cir Cardiovasc* 2015, **30**:148–158.
13. 13. Chen T, Guestrin C: **XGBoost: A Scalable Tree Boosting System.** In *Acm Sigkdd International Conference on Knowledge Discovery & Data Mining*. 2016
14. 14. O'Connor MJ, Glatz AC, Rossano JW, Shaddy RE, Ryan R, Ravishankar C, Fuller S, Mascio CE, Gaynor JW, Lin KY: **Cumulative Effect of Preoperative Risk Factors on Mortality After Pediatric Heart Transplantation.** *Ann Thorac Surg* 2018, **106**:561–566.
15. 15. Jenkins KJ, Koch Kupiec J, Owens PL, Romano PS, Geppert JJ, Gauvreau K: **Development and Validation of an Agency for Healthcare Research and Quality Indicator for Mortality After Congenital Heart Surgery Harmonized With Risk Adjustment for Congenital Heart Surgery (RACHS-1) Methodology.** *J Am Heart Assoc* 2016, **5**.
16. 16. Chu DI, Tan JM, Mattei P, Costarino AT, Rossano JW, Tasian GE: **Mortality and Morbidity after Laparoscopic Surgery in Children with and without Congenital Heart Disease.** *J Pediatr* 2017, **185**:88–93.e83.
17. 17. Hu Z, Yuan X, Rao K, Zheng Z, Hu S: **National trend in congenital heart disease mortality in China during 2003 to 2010: a population-based study.** *J Thorac Cardiovasc Surg* 2014, **148**:596–602.e591.
18. 18. Al-Radi OO, Harrell FE, Jr., Caldarone CA, McCrindle BW, Jacobs JP, Williams MG, Van Arsdell GS, Williams WG: **Case complexity scores in congenital heart surgery: a comparative study of the Aristotle Basic Complexity score and the Risk Adjustment in Congenital Heart Surgery (RACHS-1) system.** *J Thorac Cardiovasc Surg* 2007, **133**:865–875.

19. 19. Lacour-Gayet F, Clarke D, Jacobs J, Comas J, Daebritz S, Daenen W, Gaynor W, Hamilton L, Jacobs M, Maruszsewski B, et al: **The Aristotle score: a complexity-adjusted method to evaluate surgical results.** *Eur J Cardiothorac Surg* 2004, **25**:911–924.
20. 20. Jenkins KJ: **Risk adjustment for congenital heart surgery: the RACHS-1 method.** *Semin Thorac Cardiovasc Surg Pediatr Card Surg Annu* 2004, **7**:180–184.
21. 21. Joshi SS, Anthony G, Manasa D, Ashwini T, Jagadeesh AM, Borde DP, Bhat S, Manjunath CN: **Predicting mortality after congenital heart surgeries: evaluation of the Aristotle and Risk Adjustement in Congenital Heart Surgery-1 risk prediction scoring systems: a retrospective single center analysis of 1150 patients.** *Ann Card Anaesth* 2014, **17**:266–270.
22. 22. Horer J, Belli E, Roussin R, LeBret E, Ly M, Abdullah J, Marzullo R, Strbad M, Cleuziou J, Pabst von Ohain J, Lange R: **Evaluation of the Adult Congenital Heart Surgery Mortality Score at Two European Centers.** *Ann Thorac Surg* 2018, **105**:1441–1446.
23. 23. Mattos SS, Neves JR, Costa MC, Hatem TP, Luna CF: **An index for evaluating results in paediatric cardiac intensive care.** *Cardiol Young* 2006, **16**:369–377.
24. 24. Rogers L, Brown KL, Franklin RC, Ambler G, Anderson D, Barron DJ, Crowe S, English K, Stickley J, Tibby S, et al: **Improving Risk Adjustment for Mortality After Pediatric Cardiac Surgery: The UK PRAiS2 Model.** *Ann Thorac Surg* 2017, **104**:211–219.
25. 25. Russell RA, Rettiganti M, Brundage N, Jeffries HE, Gupta P: **Performance of Pediatric Risk of Mortality Score Among Critically Ill Children With Heart Disease.** *World J Pediatr Congenit Heart Surg* 2017, **8**:427–434.
26. 26. Larsen SH, Pedersen J, Jacobsen J, Johnsen SP, Hansen OK, Hjortdal V: **The RACHS-1 risk categories reflect mortality and length of stay in a Danish population of children operated for congenital heart disease.** *Eur J Cardiothorac Surg* 2005, **28**:877–881.
27. 27. Villa-Hincapie CA, Carreno-Jaimes M, Obando-Lopez CE, Camacho-Mackenzie J, Umana-Mallarino JP, Sandoval-Reyes NF: **Risk Factors for Mortality in Reoperations for Pediatric and Congenital Heart Surgery in a Developing Country.** *World J Pediatr Congenit Heart Surg* 2017, **8**:435–439.
28. 28. Almond CS, Gauvreau K, Canter CE, Rajagopal SK, Piercey GE, Singh TP: **A risk-prediction model for in-hospital mortality after heart transplantation in US children.** *Am J Transplant* 2012, **12**:1240–1248.
29. 29. Bobillo-Perez S, Sanchez-de-Toledo J: **Risk stratification models for congenital heart surgery in children: Comparative single-center study.** 2019, **14**:1066–1077.
30. 30. Singh NP, Bapi RS, Vinod PK: **Machine learning models to predict the progression from early to late stages of papillary renal cell carcinoma.** *Comput Biol Med* 2018, **100**:92–99.
31. 31. Sun Y, Du P, Lu X, Xie P, Qian Z, Fan S, Zhu Z: **Quantitative characterization of bovine serum albumin thin-films using terahertz spectroscopy and machine learning methods.** *Biomed Opt Express* 2018, **9**:2917–2929.

32. 32. Richter AN, Khoshgoftaar TM: **A review of statistical and machine learning methods for modeling cancer risk using structured clinical data.** *Artif Intell Med* 2018, **90**:1–14.
33. 33. Diller GP, Kempny A, Babu-Narayan SV, Henrichs M, Brida M, Uebing A, Lammers AE, Baumgartner H, Li W, Wort SJ, et al: **Machine learning algorithms estimating prognosis and guiding therapy in adult congenital heart disease: data from a single tertiary centre including 10 019 patients.** *Eur Heart J* 2019, **40**:1069–1077.
34. 34. Kwon JM, Kim KH, Jeon KH, Park J: **Deep learning for predicting in-hospital mortality among heart disease patients based on echocardiography.** *Echocardiography* 2019, **36**:213–218.

Figures

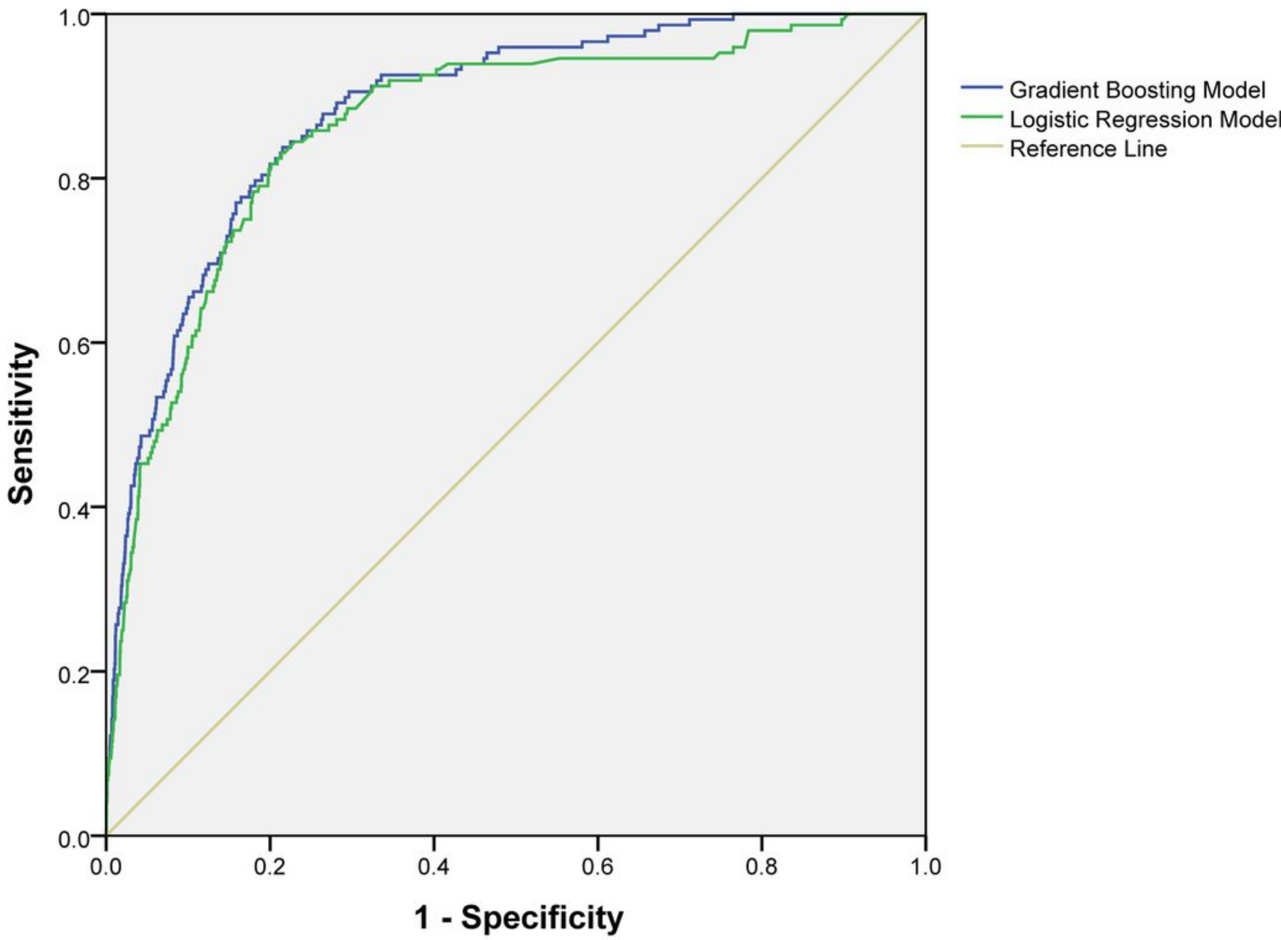

Figure 1

Comparison between prediction value (ROC curve analysis) of logistic regression vs gradient boosting model in validation set

Supplementary Files

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

- supplementaltables.docx