의학박사 학위논문

일차성 분만후출혈 응급실 환자에서 머신러닝 기반 대량수혈의 예측모델 개발

Machine learning-based prediction of massive transfusion in emergency department patients with primary postpartum hemorrhage

울산대학교 대학원

의학과

유 지 나

일차성 분만후출혈 응급실 환자에서 머신러닝 기반 대량수혈의 예측모델 개발

지도교수 손 창 환

이 논문을 의학박사학위 논문으로 제출함

2024년 1 월

울산대학교대학원

의 학 과

유지나

유지나의 의학박사학위 논문을 인준함

심사위원 안 신 인

심사위원 서동우 인

심사위원 유승목 인

심사위원 고벽성 인

심사위원 손창환 인

울산대학교 대학원

2024년 1월

Abstract

Objective: While an early massive transfusion (MT) protocol can reduce patient mortality, accurately assessing the risk of MT in emergency department (ED) patients with primary postpartum hemorrhage (PPH) presents a challenge. Therefore, the objective of this study is to identify influential factors and develop predictive models to assess the likelihood of the need for MT occurrence in ED patients with primary PPH.

Methods: Clinical features to assess variables extracted in the early and late phases of the clinical course in ED patients with primary PPH for predicting MT were retrospectively analyzed. Nine variables obtained early in the clinical course of the patients (age, sex, delivery type, initial mental status, initial vital signs, shock index, and lactate) were categorized as 'early features'. In addition, 'all features' were analyzed, including all variables obtained in the late phase: hemoglobin, hematocrit, platelet counts, blood urea nitrogen, creatinine, prothrombin time, fibrinogen, d-dimer, fibrin degradation product, and albumin. Seven machine learning algorithm models have been developed using early and all features, respectively: Support Vector Machine with Polynomial Kernel (SVM poly), Support Vector Machine with Radial Basis Function Kernel (SVM radial), K-Nearest Neighbors (KNN), Extreme Gradient Boosting (XGBoost), Logistic Regression (Logistic), Random Forest, and Decision Tree.

Results: Out of the 612 patients, 101 (33.4%) required MT.

When all features were analyzed, the Random Forest model showed the highest performance. The area under the curve values, accuracy, sensitivity, and specificity of the prediction model for MT risk prediction were 0.885, 89%, 74%, and 95%, respectively.

Conclusion: This study provides a relatively accurate and applicable identification of risk

factors associated with MT that may be useful for improving hemorrhagic shock management in ED patients with PPH patients.

Contents

Lists of Tables

Lists of Figures

Introduction

Maternal mortality refers to the death of a pregnant woman during pregnancy or within 42 days of the termination of pregnancy. Maternal mortality ratio is the number of maternal deaths per 100,000 live births. According to a 2014 report by the World Health Organization, the global maternal mortality rate has steadily decreased from 283 to 209 between 1990 and 2013, while South Korea also experienced a similar decrease from 21 to 12 in 1990. [1]

Uncontrolled bleeding after delivery, known as postpartum hemorrhage (PPH), is a significant contributor to maternal mortality. It accounts for nearly 20% of all maternal deaths worldwide, with an estimated 125,000 deaths per year [2-4]. The United States has one of the highest maternal mortality rates among developed countries, with approximately 11% of all maternal deaths associated with PPH. PPH, hypertensive disorders of pregnancy, and obstetric embolism are the three main causes of maternal death.

PPH is considered a leading and manageable cause of maternal mortality and morbidity worldwide. PPH accounts for 19.7% of maternal deaths in developed regions, compared to 8% in developing countries [3].

PPH is traditionally defined as blood loss of 500 mL or more in the first 24 hours after delivery. However, previous studies have shown that physicians' estimates of bleeding amounts during delivery were unreliable, and clinical management could be delayed if decisions are based on these estimates [5]. In clinical practice, it is crucial to promptly and objectively identify patients at risk of severe hemorrhagic shock that may require massive transfusion (MT) in order to optimize patient outcomes.

Initial factors, such as initial lactate and shock index, have been shown to predict the need for MT in patients with primary PPH [6,7]. Furthermore, predictive models that combine coagulopathy, vital signs, and trauma vectors have been proposed for other trauma patients [8-10]. However, the practical application of a single effective prediction model is limited for MT prediction. Machine learning is a subfield of artificial intelligence that enables algorithms to improve their performance on specific tasks using empirical data.

Therefore, the objective of this study is to identify influential factors and construct predictive models to assess the likelihood of the need for MT in ED patients with primary PPH.

Methods

Study design and data source

A retrospective cohort study was conducted on 612 patients with primary PPH who were referred to the ED of Asan Medical Center, tertiary hospital in Seoul, South Korea, between January 2004 and August 2023. This study was approved by the institutional review board of our institution and the requirement for written informed consent was waived due to the retrospective nature of the study.

The study patients were initially identified through a hospital computer database system using a hospital discharge diagnosis of ‘Postpartum hemorrhage’. Primary PPH was defined as hemorrhage requiring transfusion or fluid resuscitation within the first 24 hours after delivery.

None of the patients included in this study delivered at our hospital; instead, they were referred to our ED for the evaluation and management of PPH after delivery.

The medical records, including maternal demographics, clinical findings, and outcomes, were reviewed. Initial vital signs were used to calculate the shock index, which is defined as the pulse rate divided by the systolic blood pressure, and the modified shock index, which is defined as the pulse rate divided by the mean arterial pressure. Laboratory tests were also performed upon arrival at the ED and investigated.

The primary outcome of this study was the need for MT. MT was defined as the transfusion of 10 units or more of packed red blood cells within the initial 24 hours after the onset of PPH. Both the amount of blood transfused before arrival at the ED and the amount of blood transfused after arrival at the ED were measured to determine the need for MT. The goal of the blood transfusion decision was to improve the patient's symptoms, hemodynamics, and

laboratory results as determined by the physician. Secondary outcomes included the performance of emergency hysterectomy, embolization for bleeding control, ICU admission, length of hospital stay, and mortality.

Statistic technique and prediction model

The framework for developing each predictive machine learning model is shown in Figure 1. Patients with primary PPH were divided into a training or testing cohort using data mining. Univariate and multivariate analyses were performed to confirm and validate the classification of each cohort. We initially performed a statistical analysis to evaluate the significance of each variable investigated for all patients with primary PPH. Univariate analyses were performed on each variable and MT to determine the mean, standard deviation, median, and the proportion at the 25th and 75th interquartile ranges. A Student's t-test was used to analyze significant differences in the quantitative parameters, while a Fisher's exact test was used for the qualitative parameters. Statistical significance was set at $p < 0.05$. Multivariate analyses were performed to determine predictors for MT and to select significant variables for the prediction model.

For model training, the entire data set was randomly divided into a training set (75% for model development and optimization) and a validation set (25% for model testing). In each set, random down-sampling was performed to enable a 1:1 analysis between the MT group and the non-MT group. Linear regression imputations were used to address missing values. The predictive variables were divided into two groups: "Early features", which included 9 variables available immediately upon arrival, and "All features", which included 19 variables, including laboratory test results and early features.

1) Early features: age, parity, delivery type, initial mental status, initial vital signs (systolic blood pressure, diastolic blood pressure, pulse rate), shock index, and lactate

2) All features: Early features + hemoglobin, hematocrit, platelet count, BUN (blood urea nitrogen), creatinine, prothrombin time, fibrinogen, D-dimer, FDP (fibrin degradation product), and albumin.

Hyperparameter tuning was performed to find the best fit for each model using a coarse-grained grid search with repeated random subsampling for cross-validation. The 75:25 train-validation split was used for model validation in both the early features cohort and the all features cohort (Figure 2).

The ratio of MT groups to non-MT groups was maintained in both the training set and validation set to match the original dataset. The average performance was calculated using ten-fold cross-validation as 9:1. We used the training set to determine the parameters of the prediction models and evaluated their performance using the “validation set”.

To predict the need for MT in ED patients with primary PPH, seven machine learning models were used. These models used both the established 9 variables of early features and 19 variables of all feature sets. The models included Support Vector Machine with Polynomial Kernel (SVM poly), Support Vector Machine with Radial Basis Function Kernel (SVM radial), K-Nearest Neighbors (KNN), Extreme Gradient Boosting (XGBoost), Logistic Regression (Logistic), Random Forest, and Decision Tree.

For performance evaluation, the accuracy of repeated models with narrow variations in efficiency was assessed for each machine learning algorithm. The estimated accuracy for the

classification model tuning results is shown in Figure 3, including both preterm features and all features.

The best performing models were selected for the predictive model, and a validation set was applied to derive the predicted outcome results. Predictive models were developed and validated using both early features and all features with seven machine learning models.

Sensitivity, specificity, positive predictive value, negative predictive value, and F1 score were assessed in each model. Discrimination was assessed using the receiver operating characteristic curve to calculate area under the curve for each machine learning analysis. Statistical analyses and machine learning were implemented using R statistical software (Version 3.6.1).

Results

Clinical Characteristics of the Study Patients

Out of 612 patients with primary PPH, 169 patients (27.6%) experienced MT. The characteristics of ED patients with primary PPH in the MT group and non-MT group are described in Table 1 and Figure 2.

The median age of ED patients with primary PPH was higher in the MT group (33 [30-36] vs. 32 [29-35], P=0.028). Of all the patients, 355 (59.8%) were primiparous and 246 (40.2%) were multiparous. There was no difference in the incidence of MT between the two groups.

Out of 612 patients, 407 (66.5%) underwent vaginal delivery, and 205 (33.5%) underwent Caesarean delivery. The delivery type was not statistically significant in relation to the

occurrence of MT. Thirty-one (5.1%) patients showed a change in mental status upon arrival at the ED from the hospital of delivery. More patients with initial consciousness problems required compared to others [(23 (74.2%) vs. 8 (25.8%), P<0.05].

The average systolic blood pressure for all patients was 112 ± 23 mmHg. In the MT group, it was lower at 102 ± 26 mmHg compared to 116 ± 20 mmHg in the non-MT group (P<0.05). Similarly, the average diastolic blood pressure was 71 ± 17 mmHg for all patients, with 65 ± 22 mmHg in the MT group and 72 ± 17 mmHg in the non-MT group (P<0.05). Pulse rate was significantly higher in the MT group than in the non-MT group (114 [103-127] vs. 90 [81-102], P<0.05). There was no significant difference in body temperature between the two groups (P=0.117)

The shock index was significantly higher in the MT group, with median values of 1.23 compared to 0.8 in the non-MT group (P<0.05).

In the laboratory test results, the MT group showed significantly lower levels of hemoglobin (8.5 ± 2.9 vs. 10.1 ± 2.0 g/dL, P<0.05) and hematocrit (26.0 ± 8.5 vs. 30.5 ± 5.7%, P<0.05) compared to the non-MT group. The platelet count was significantly lower in the MT group (107 [91-142] vs. 162 [134-199] x10^3/μL, P<0.05). Blood urea nitrogen and creatinine levels were significantly higher in the MT group compared to the non-MT group (P<0.05 for both).

The MT group had significantly lower albumin levels (1.4 [1.2-1.8] vs. 2.1 [1.9-2.3] mg/dL, P<0.05). Elevated lactate levels were more prominent in the MT group (3.1 [1.9-4.3] vs. 2.1 [1.6-2.9], P<0.05). The MT group showd a longer prothrombin time, INR (1.4 [1.2-2.0] vs. 1.1 [1.0-1.2], P<0.05), and decreased fibrinogen levels (110.5 ± 83.3 vs. 246.1 ± 106.2, P<0.05). D-dimer levels were significantly lower in the MT group (58.2 [8.4-35.5] vs. 13.1 [6.8-35.2], P<0.05). FDP levels were significantly higher in the MT group (112.5 [24.2-

120.0] vs. 44.4 [21.7-120.0], P<0.05). No significant difference was observed in the white blood cell counts and C-reactive protein levels between the MT group and non-MTgroup (Table 2).

In terms of outcomes, a higer percentage of patients in the non-MT group underwent embolization compared to the MT group (54.0% vs. 46.0%, P<0.05). Conversely, a higher percentage of patients in the MT group underwent postpartum hysterectomy (83.3% vs. 16.7%, P<0.05).

All mortality cases were observed in the MT group (100% vs. 0%, P<0.05), and the length of hospital stay for patients in the MT group was longer, with a median of 4 days [3-7 days], compared to 2 days [1-3 days] for patients in the non-MT group (P<0.05) (Table 3).

The training set and validation set were divided into 458 and 154 samples, respectively, using a 75:25 ratio. MT occurred in 126 in the training set, and 111 in the validation set (Table 4 and Table 5). Nineteen features were compared between the MT group and non-MT group in bothe the training set and validation set.

The median age of ED patients with PPH was higher in the MT group (33 [31-36] vs. 32 [29-35], P=0.017) in the training set. On the other hand, there was no significant difference in the median ages in the validation set. In both sets, there was no significant difference in the occurrence of MT based on parity and delivery type.

In the early feature variables, the average systolic blood pressure and diastolic blood pressure were lower in the MT group than in the non-MT group in the training set (110 ± 27 vs. 116 ± 20, P<0.05) and validation set (94 ± 23 vs. 117 ± 22, P<0.05). The average diastolic blood pressure was significantly different between the MT group and the non-MT group, with 65 ±

23 mmHg vs. 72 ± 16 mmHg in the validation set, and 59 ± 19 mmHg vs. 72 ± 18 mmHg in the validation set (P<0.05). The pulse rate was higher in the MT group than in the non-MT group in both the training set and validation set (115 [97-128] and 113 [94-139] compared to 92 [81-105] and 92 [79-104], P<0.05). The shock index was significantly higher in the MT group compared to the non-MT group in both the training set and validation set (1.12 [0.89-1.45] and 1.12 [0.90-1.54] compared to 0.79 [0.67-0.96] and 0.78 [0.67-0.95], p<0.05). Lactate levels were higher in the MT group (3.9 [2.5-5.3] and 3.4 [2.4-5.9] vs. 2.2 [1.7-3.2] and 2.1 [1.6-2.9], P<0.05) in both sets.

In the late feature variables, the MT group showed a significantly lower mean hemoglobin level (8.5 ± 2.8 vs. 10.1 ± 2.0 g/dL, P<0.05) and hematocrit (26.2 ± 8.2 vs. 30.6 ± 5.7, P<0.05) compared to the non-MT group in the training set. Similarly, in the validation set, the hemoglobin levels (8.5 ± 3.1 vs. 10.3 ± 2.1 g/dL, P<0.05) and hematocrit levels (26.1 ± 9.2 vs. 31.2 ± 5.8, P<0.05) lower than those in the non-MT group. The platelet counts were lower in the MT group (109 [90-140] and 106 [88-146] vs. 162 [134-199] and 162 [131-196], P<0.05 in both) in the training and validation sets. In the MT group, BUN and creatinine levels were higher compared to the non-MT group in both the training set and validation set, although they were within the normal range. (9.0 [7.0-11.0] and 9.0 [7.0-12.3] vs. 7.0 [6.0-9.0] and 6.5 [6.0-9.0], 0.66 [0.56-0.89] and 0.75 [0.51-0.81] vs. 0.55 [0.49-0.80] and 0.54 [0.47-0.65], P<0.05 for both). The MT group also had significantly lower albumin levels (1.5 [1.1-1.8] and 1.5 [1.1-1.8] vs. 2.1 [1.7-2.3] and 2.1 [1.8-2.3] mg/dL, P<0.05 for both) in both sets. PT INR (1.5 [1.2-2.5] and 2.3 [1.4-2.3] vs. 1.1 [1.0-1.3] and 1.2 [1.1-1.5], P<0.05 for both), D-dimer (35.2 [8.0-71.5] and 35.4 [17.0-53.1] vs. 12.8 [5.3-35.2] and 10.2 [4.4-34.6], P<0.05), and fibrinogen levels (113.5 ± 83.4 and 100.2 ± 74.2 vs. 244.2 ± 106.8 and 259.1 ±

112.0, P<0.05) were lower in the MT group in both sets. FDP also showed significantly higher levels in the MT group than in the non-MT group in both the training set (52.1 [19.7-80.3] vs. 36.4 [15.2-80.3], P=0.038) and the validation set (83.4 [41.3-125.0] and 32.2 [15.2-53.0], P<0.05).

Multivariate analysis was conducted to identify predictors of MT. The analysis used a backward stepwise logistic regression model for the entire study patients, as well as for the training set and validation set (Table 6, Table 7, and Table 8).

In analysis of all study patients, the initial shock index (OR 18.093 [95% CI 5.658-57.864], P<0.05) , INR (OR 7.103 [95% CI 2.998-16.831], P<0.05), Albumin (OR 0.343 [95% CI 0.124-0.952], P<0.05), lactate (OR 1.217 [95% CI 1.028-1.439], P<0.05) and fibrinogen (OR 0.990 [95% CI 0.986-1.002], P<0.05) were significant predictors for MT.

In the analysis of the patient group in the training set, the initial shock index (OR 10.186 [95% CI 2.862-36.259], p<0.05) was also a potent predictor for MT. Fibrinogen also showed an association with MT (OR 0.990 [95% CI 0.985-1.002], P<0.05).

In the analysis of the patient group in the validation set, only the initial shock index showed a strong predictor for MT (OR 135.061 [95% CI 5.673-3215.326], P<0.05).

Model Performance

To identify high-performance machine learning models for predicting MT in ED patients with primary PPH, seven machine learning algorithms were implemented using 9 early features and 19 all features.

The confusion matrix generated by seven machine learning algorithms using preterm features when running the prediction model with the validation set is shown in Figure 4. Figure 5 shows the confusion matrix generated using all 19 features with the validation set.

In demonstrating the relative importance of variables for predicting MT in a machine learning model, Random Forest, XGBoost, and Logistic Regression models were compared using both early and all features. Lactate and shock index were commonly identified as potent variables for predicting MT in patients with primary PPH when using a machine learning algorithm with early features (Figure 6). When analyzing machine learning with Random Forest, XGBoost, and Logistic Regression models using all features, it was found that Prothrombin time, fibrinogen, and hemoglobin were important variables (Figure 7).

In each set of variables, the performance parameters of machine learning algorithms were derived using both early and all features.

Among the machine learning algorithms that use early features, the SVM radial model demonstrated the highest ROC of 0.783 and the highest accuracy, reaching 0.746. Table 9 and Figure 8 present a comparison of the area under the curve for each machine learning analysis using early features in the training and validation sets.

In the implementation of the model using 22 features, the Random Forest model achieves the highest ROC of 0.884 and the highest accuracy, 0.876. The comparison includes the area under the curve for each machine learning analysis with all predictors in the training and validation sets (Table 10 and Figure 9).

When using only early features, the SVM radial model showed better performance in most metrics. When considering all features, the Random Forest model was one of the top

contenders for predicting MT in patients with primary PPH. Overall, while the SVM radial consistently performed well in both scenarios, it is important to note that other models may also demonstrate commendable performance depending on the predictors used.

Discussion

This study identified influential factors and constructed predictive models to assess the likelihood of the need for MT in ED patients with primary PPH. A prediction model based on the SVM radial and Random Forest has been successfully constructed, demonstrating superior performance compared to other machine learning models and logistic regression.

This study is the first to develop a machine learning model for predicting the need for MT in ED patients with primary PPH. PPH is a leading cause of preventable death. Early identification of ED patients with primary PPH who require MT is crucial to maximize the benefits of early transfusion administration.

MT is an adverse outcome associated with poor prognosis in various trauma and medical conditions involving hemorrhage. Many traditional statistical methods have been used to predict the need for MT and evaluate the risk of MT. Various predictors have been suggested.

Akaraborworn et al. [11,12] predicted the need for MT in trauma patients admitted to the surgical intensive care unit by using the Sequential Organ Failure Assessment (SOFA) score, intraoperative blood loss, age, base excess, and heart rate. Coagulopathy on thromboelastography was suggested as a predictor of the need for MT in trauma patients [13,14].

Lactate > 4 mmol/L has also been suggested as a predictor of the need for MT in hemodynamically stable trauma patients [15]. Narrow pulse pressure, shock index, and

modified shock index are also recognized as predictors of the need for MT and mortality in trauma patients [10,16-18].

There have been fewer studies on PPH compared to trauma. The study already shown that MT in patients with PPH can be predicted by lactate levels, coagulopathy, and shock index [6,7,19,20]. Similarly, the ‘obstetric shock index’, which measures the shock index during the peripartum period, proved to be useful in estimating blood loss in PPH [21]. Recently, coagulopathy, such as hypofibrinogenemia in viscoelastic hemostatic assays, has been associated with the occurrence and treatment of massive PPH [22,23].

Machine learning is a subfield of artificial intelligence that enables algorithms to improve their performance on specific tasks by using empirical data. The use of machine learning has resulted in the development of prognostic models that can be effectively applied in clinical settings. These models analyze multiple interconnected variables. This study used various machine learning algorithms to identify the risk factors associated with MT. A prediction model was developed using the training dataset and validated using the testing dataset. Predictive models can enable healthcare practitioners to quickly assess a patient's condition, allowing for timely medical interventions. The study has the potential to improve patient safety and reduce complications associated with primary PPH.

Despite our promising findings, there are still limitations in our study. First, the study was conducted in a specific clinical situation where MT occurred in a relatively rare and fatal clinical course of PPH. In addition, although MT occurred in 27% of patients, a 1:1 down-sampling for machine learning was performed, which may have resulted in the loss of important information. Second, the limitations of a retrospective study resulted in significant missing data for several variables. Statistically imputed missing data can reduce the statistical

power of a study and produce biased estimates, leading to invalid conclusions. The counts of missing FDP, fibrinogen, lactate, and D-dimer were 123 (20.1%), 102 (16.7%), 97 (15.8%), and 87 (14.2%), respectively. Third, the small sample size of this single-center study is a significant limitation. External validation should be considered to obtain clinical results and avoid the potential risk of overfitting. The MT prediction model has not yet been validated in a prospective study. In the future, it is necessary to conduct multicenter and larger-scale research. To develop a machine learning prediction model with higher sensitivity and specificity. This model can be used to accurately assess the risk of MT.

Conclusions

It has been demonstrated that machine learning prediction methods can be used to develop accurate prediction models for MT in patients with primary PPH. A prediction model was successfully developed using Random Forest, which may outperform the traditional logistic regression method. These predictive models are unique in their kind due to our expertise. We are working to improve accuracy by learning from ongoing observational data and are committed to integrating it into future clinical workflows.

Reference

1. Kassebaum NJ, Bertozzi-Villa A, Coggeshall MS, Shackelford KA, Steiner C, Heuton KR, et al. Global, regional, and national levels and causes of maternal mortality during 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013. Lancet 2014;384:980-1004.
2. Selo-Ojeme DO. Primary postpartum haemorrhage. J Obstet Gynaecol 2002;22:463-9.
3. Say L, Chou D, Gemmill A, Tuncalp O, Moller AB, Daniels J, et al. Global causes of maternal death: a WHO systematic analysis. Lancet Glob Health 2014;2:e323-33.
4. Fink DA, Kilday D, Cao Z, Larson K, Smith A, Lipkin C, et al. Trends in Maternal Mortality and Severe Maternal Morbidity During Delivery-Related Hospitalizations in the United States, 2008 to 2021. JAMA Netw Open 2023;6:e2317641.
5. Ashburn JC, Harrison T, Ham JJ, Strote J. Emergency physician estimation of blood loss. West J Emerg Med 2012;13:376-9.
6. Sohn CH, Kim YJ, Seo DW, Won HS, Shim JY, Lim KS, et al. Blood lactate concentration and shock index associated with massive transfusion in emergency department patients with primary postpartum haemorrhage. Br J Anaesth 2018;121:378-83.
7. Sohn CH, Kim WY, Kim SR, Seo DW, Ryoo SM, Lee YS, et al. An increase in initial shock index is associated with the requirement for massive transfusion in emergency department patients with primary postpartum hemorrhage. Shock 2013;40:101-5.
8. Strickland M, Nguyen A, Wu S, Suen SC, Mu Y, Del Rio Cuervo J, et al. Assessment of Machine Learning Methods to Predict Massive Blood Transfusion in Trauma.

World J Surg 2023;47:2340-6.

9. D'Souza K, Norman M, Greene A, Finney CJF, Yan MTS, Trudeau JD, et al. Prediction of massive transfusion with the Revised Assessment of Bleeding and Transfusion (RABT) score at Canadian level I trauma centers. Injury 2023;54:19-24.

10. Carsetti A, Antolini R, Casarotta E, Damiani E, Gasparri F, Marini B, et al. Shock index as predictor of massive transfusion and mortality in patients with trauma: a systematic review and meta-analysis. Crit Care 2023;27:85.

11. Akaraborworn O, Chaiwat O, Chatmongkolchart S, Kitsiripant C, Chittawatanarat K, Morakul S, et al. Prediction of massive transfusion in trauma patients in the surgical intensive care units (THAI-SICU study). Chin J Traumatol 2019;22:219-22.

12. Akaraborworn O, Siribumrungwong B, Sangthong B, Thongkhao K. Massive Blood Transfusion for Trauma Score to Predict Massive Blood Transfusion in Trauma. Crit Care Res Pract 2021;2021:3165390.

13. Cohen J, Scorer T, Wright Z, Stewart IJ, Sosnov J, Pidcoke H, et al. A prospective evaluation of thromboelastometry (ROTEM) to identify acute traumatic coagulopathy and predict massive transfusion in military trauma patients in Afghanistan. Transfusion 2019;59:1601-7.

14. Barrett CD, Moore HB, Vigneshwar N, Dhara S, Chandler J, Chapman MP, et al. Plasmin thrombelastography rapidly identifies trauma patients at risk for massive transfusion, mortality, and hyperfibrinolysis: A diagnostic tool to resolve an international debate on tranexamic acid? J Trauma Acute Care Surg 2020;89:991-8.

15. Brooke M, Yeung L, Miraflor E, Garcia A, Victorino GP. Lactate predicts massive transfusion in hemodynamically normal patients. J Surg Res 2016;204:139-44.

16. David JS, Voiglio EJ, Cesareo E, Vassal O, Decullier E, Gueugniaud PY, et al.

Prehospital parameters can help to predict coagulopathy and massive transfusion in trauma patients. Vox Sang 2017;112:557-66.

17. Wang IJ, Bae BK, Park SW, Cho YM, Lee DS, Min MK, et al. Pre-hospital modified shock index for prediction of massive transfusion and mortality in trauma patients. Am J Emerg Med 2020;38:187-90.

18. Warren J, Moazzez A, Chong V, Putnam B, Neville A, Singer G, et al. Narrowed pulse pressure predicts massive transfusion and emergent operative intervention following penetrating trauma. Am J Surg 2019;218:1185-8.

19. Sohn CH, Kim SR, Kim YJ, Seo DW, Ahn S, Lee YS, et al. Disseminated Intravascular Coagulation in Emergency Department Patients With Primary Postpartum Hemorrhage. Shock 2017;48:329-32.

20. Kong T, Lee HS, Jeon SY, You JS, Lee JW, Chung HS, et al. Delta neutrophil index and shock index can stratify risk for the requirement for massive transfusion in patients with primary postpartum hemorrhage in the emergency department. PLoS One 2021;16:e0258619.

21. Le Bas A, Chandraharan E, Addei A, Arulkumaran S. Use of the "obstetric shock index" as an adjunct in identifying significant blood loss in patients with massive postpartum hemorrhage. Int J Gynaecol Obstet 2014;124:253-5.

22. Bell SF, Roberts TCD, Freyer Martins Pereira J, De Lloyd L, Amir Z, James D, et al. The sensitivity and specificity of rotational thromboelastometry (ROTEM) to detect coagulopathy during moderate and severe postpartum haemorrhage: a prospective observational study. Int J Obstet Anesth 2022;49:103238.

23. Bell SF, Collis RE, Collins PW. Comparison of haematological indices and transfusion management in severe and massive postpartum haemorrhage: analysis of a

two-year national prospective observational study. Int J Obstet Anesth 2022;50:103547.

국문 요약

연구 배경 및 목적: 초기에 대량 수혈 프로토콜로 환자 사망률을 줄일 수 있음에도, 임상 초기에서 일차 분만후출혈 환자에서 대량수혈의 필요성을 정확하게 파악하는 것은 여전히 주요 과제로 남아 있다.

연구 방법: 일차성 분만후출혈 환자에서 임상 경과 초기 및 말기에 추출한 변수를 이용하여 대량수혈 예측을 위한 임상적 특징을 후향적으로 분석하였다. 환자의 임상 경과 초기에 얻은 9 가지 초기 변수(나이, 성별, 분만 유형, 초기 의식 상태, 초기 활력 징후, 쇼크 지수, 젖산염)과 후기에 얻어진 헤모글로빈, 헤마토크릿, 혈소판 수, 혈중 요소 질소, 크레아티닌, 프로트롬빈 시간, 피브리노겐, d-이합체, 피브린 분해 산물, 알부민을 모두 포함한 19 가지 변수를 각각 나누어 분석하였다. 분석에는 다항식 커널을 이용한 서포트 벡터 머신(SVM poly), 방사형 기저함수 커널을 이용한 서포트 벡터 머신(SVM radial), K-최근접 이웃(KNN), 극한 경사 부스팅(XGBoost), 로지스틱 회귀, 랜덤 포레스트, 의사결정 트리 등 7 가지 머신러닝 알고리즘 모델을 각각 초기 및 전체 변수들에 대해서 각각 시행되었다.

결과 및 고찰: 612 명의 환자 중 101 명(33.4%)의 환자에서 대량수혈이 필요하였다. 355 명(59.8%)의 환자가 초산이었으며, 407 명(66.5%)의 환자가 질식분만 후 일차성 분만후출혈이 발생하였다. 모든 변수를 분석에 사용하였을 때 랜덤 포레스트 모델에서 가장 높은 성능이 관찰되었으며, 대량수혈 위험도

예측 모델의 AUC, 정확도, 민감도, 특이도 값은 각각 89%, 74%, 95%, 0.885 로 나타났다. 분석 결과, 젖산염, 쇼크 지수, 초기 의식 상태가 초기 대량수혈의 가장 중요한 위험 요인으로 나타났다. 혈소판 시간, 초기 혈청 단백질 또는 알부민 수치, 헤모글로빈도 전체 기간의 대량수혈 예측 모델에 중요한 요인으로 나타났다.

결론: 본 연구는 일차성 분만후출혈 응급실 환자에서 비교적 정확하고 임상적으로 적용가능한 대량수혈 예측모델을 개발함으로써 대량수혈의 위험인자와 출혈성 쇼크 관리가 이루어 질 수 있도록 하였다.

Table 1. Baseline and clinical characteristics of the study patients

Characteristics	All patients (n=612)	MT group (n=169)	Non-MT group (n=443)	*P* value
Age, years	32 [30-35]	33 [30-36]	32 [29-35]	0.028
Parity				0.703
Primipara	355 (59.8)	99 (27.0)	267 (73.0)	
Multipara	246 (40.2)	70 (28.5)	176 (71.5)	
Delivery type				0.400
Vaginal delivery	407 (66.5)	108 (26.5)	299 (73.5)	
Caesarean section	205 (33.5)	61 (29.8)	144 (70.2)	
Altered mental status	31 (5.1)	23 (74.2)	8 (25.8)	<0.001
Initial vital signs				
Systolic blood pressure, mmHg	112 ± 23	102 ± 26	116 ± 20	<0.001
Diastolic blood pressure, mmHg	71 ± 17	65 ± 22	72 ± 17	<0.001
Respiratory rate, rates/min	20 [18-20]	20 [18-22]	20 [10-20]	<0.001
Pulse rate, beats/min	97 [83-115]	114 [103-127]	90 [81-102]	<0.001
Body temperature, ℃	37.4 ± 0.8	37.1 ± 0.9	37.3 ± 0.9	0.117
Shock index	0.82 [0.70-1.04]	1.23 [0.83-1.59]	0.8 [0.67-0.96]	<0.001

Values are expressed as the mean ± standard deviation, median [interquartile range], or number (%).

Table 2. Laboratory findings of the study patients

Characteristics	All patients (n=612)	MT group (n=169)	Non-MT group (n=443)	*P* value
White blood cells, $\times 10^3/\mu L$	17.9 ± 5.5	17.9 ± 6.5	17.9 ± 5.4	0.934
Hemoglobin, g/dL	9.9 ± 2.2	8.5 ± 2.9	10.1 ± 2.0	<0.001
Hematocrit, %	30.3 ± 6.1	26.0 ± 8.5	30.5 ± 5.7	<0.001
Platelets, $\times 10^3/\mu L$	152 [118-191]	107 [91-142]	162 [134-199]	<0.001
Blood urea nitrogen, mg/dL	7.5 [6.0-10.0]	9 [6-12]	7 [6-10]	<0.001
Creatinine, mg/dL	0.54 [0.46-0.63]	0.77 [0.61-1.04]	0.62 [0.54-0.77]	<0.001
Albumin, mg/dL	2.2 [1.6-3.1]	1.4 [1.2-1.8]	2.1 [1.9-2.3]	<0.001
C-reactive protein, mg/dL	0.58 [0.26-1.16]	0.61 [0.21-1.15]	0.58 [0.27-1.17]	0.07
Lactate, mmol/L	2.2 [1.6-3.1]	3.1 [1.9-4.3]	2.1 [1.6-2.9]	<0.001
Prothrombin time, INR	1.1 [1.0-1.2]	1.4 [1.2-2.0]	1.1 [1.0-1.2]	<0.001
D-dimer, μg/mL	15.2 [7.5-35.5]	35.2 [8.4-35.5]	13.1 [6.8-35.2]	<0.001
Fibrinogen, mg/dL	229.8 ± 114.0	110.5 ± 83.3	246.1 ± 106.2	<0.001
FDP, mg/dL	52.6 [22.0-120.0]	112.5 [24.2-120.0]	44.4 [21.7-120.0]	<0.001

Values are expressed as the mean ± standard deviation or median [interquartile range].
INR, international normalized ratio; FDP, fibrin degradation product.

Table 3. Outcomes of the study patients

Characteristics	All patients (n=612)	MT group (n=169)	Non-MT group (n=443)	*P* value
Embolization	309 (50.5)	142 (46.0)	167 (54.0)	<0.001

Hysterectomy	6 (1.0)	5 (83.3)	1 (16.7)	<0.001
ICU admission	51 (8.3)	39 (76.5)	12 (23.5)	<0.001
Mortality	5 (0.8)	5 (100)	0 (0)	<0.001
Hospital length of stay	2 [1-4]	4 [3-7]	2 [1-3]	<0.001

Values are expressed as median [interquartile range], or number (%).

Table 4. Variables of the study patients in the training set

Characteristics	All patients (n=458)	MT group (n=126)	Non-MT group (n=332)	*P* value
Age, years	32 [30-35]	33 [31-36]	32 [29-35]	0.017
Parity				0.449
Primipara	271 (59.2)	71 (26.2)	200 (73.8)	
Multipara	187 (40.8)	55 (29.4)	135 (70.6)	
Delivery type				0.984
Vaginal delivery	305 (66.6)	84 (27.5)	221 (72.5)	
Caesarean section	153 (33.4)	42 (24.3)	111 (72.5)	
Altered mental status	25 (5.1)	17 (68.0)	8 (32.0)	<0.001
Initial vital signs				

Systolic blood pressure, mmHg	112 ± 23	102 ± 27	116 ± 20	<0.001
Diastolic blood pressure, mmHg	69 ± 19	65 ± 23	72 ± 16	<0.001
Pulse rate, beats/min	97 [83-116]	115 [97-128]	92 [81-105]	<0.001
Shock index	0.86 [0.73-1.08]	1.12 [0.89-1.45]	0.79 [0.67-0.96]	<0.001
Lactate, mmol/L	2.5 [1.8-3.8]	3.9 [2.5-5.3]	2.2 [1.7-3.2]	<0.001
Hemoglobin, g/dL	9.7 ± 2.4	8.5 ± 2.8	10.1 ± 2.0	<0.001
Hematocrit, %	29.7 ± 6.7	26.2 ± 8.2	30.6 ± 5.7	<0.001
Platelets, $\times 10^3/\mu L$	147 [114-188]	109 [90-140]	162 [131-196]	<0.001
Blood urea nitrogen, mg/dL	7.0 [6.0-10.0]	9.0 [7.0-11.0]	7.0 [6.0-9.0]	<0.001
Creatinine, mg/dL	0.57 [0.49-0.70]	0.66 [0.56-0.89]	0.55 [0.49-0.80]	<0.001
Albumin, mg/dL	1.9 [1.6-2.2]	1.5 [1.1-1.8]	2.1 [1.7-2.3]	<0.001
Prothrombin time, INR	1.1 [1.0-1.3]	1.5 [1.2-2.5]	1.1 [1.0-1.3]	<0.001
D-dimer, μg/mL	15.4 [5.9-35.5]	35.2 [8.0-71.5]	12.8 [5.3-35.2]	<0.001
Fibrinogen, mg/dL	208.2 ± 116.6	113.5 ± 83.4	244.2 ± 106.8	<0.001
FDP, mg/dL	39.8 [16.0-104.4]	52.1 [19.7-80.3]	36.4 [15.2-80.3]	0.038

Values are expressed as the mean ± standard deviation, median [interquartile range], or number (%).

INR, international normalized ratio; FDP, fibrin degradation product.

Table 5. Baseline and clinical characteristics of the study patients in the validation set

Characteristics	All patients (n=154)	MT group (n=111)	Non-MT group (n=43)	*P* value
Age, years	32 [30-35]	33 [30-36]	32 [30-35]	0.840

Parity				0.561
Primipara	91 (59.1)	27 (29.7)	64 (70.3)	
Multipara	63 (40.9)	16 (25.4)	47 (74.6)	
Delivery type				0.609
Vaginal delivery	110 (71.4)	32 (29.1)	78 (70.9)	
Caesarean section	44 (28.6)	11 (25.0)	33 (75.0)	
Altered mental status	6 (3.9)	6 (100)	0 (0)	<0.001
Initial vital signs				
Systolic blood pressure, mmHg	112 ± 25	94 ± 23	117 ± 22	<0.001
Diastolic blood pressure, mmHg	70 ± 20	59 ± 19	72 ± 18	<0.001
Pulse rate, beats/min	96 [82-116]	113 [94-139]	92 [79-104]	<0.001
Shock index	0.82 [0.72-1.18]	1.12 [0.90-1.54]	0.78 [0.67-0.95]	<0.001
Lactate, mmol/L	2.5 [1.8-3.8]	3.4 [2.4-5.9]	2.2 [1.6-3.1]	<0.001
Hemoglobin, g/Dl	9.7 ± 2.5	8.5 ± 3.1	10.3 ± 2.1	<0.001
Hematocrit, %	29.4 ± 7.3	26.1 ± 9.2	31.2 ± 5.8	<0.001
Platelets, $\times 10^3/\mu L$	147 [110-185]	106 [88-146]	158 [133-198]	<0.001
Blood urea nitrogen, mg/dL	7.5 [6.0-10.0]	9.0 [7.0-12.3]	6.5 [6.0-9.0]	<0.001
Creatinine, mg/dL	0.57 [0.49-0.70]	0.75 [0.51-0.81]	0.54 [0.47-0.65]	<0.001
Albumin, mg/dL	1.9 [1.0-2.3]	1.5 [1.1-1.8]	2.1 [1.8-2.3]	<0.001
Prothrombin time, INR	1.1 [1.0-1.4]	2.3 [1.4-2.3]	1.2 [1.1-1.5]	<0.001

D-dimer	14.4 [4.5-35.5]	35.4 [17.0-53.1]	10.2 [4.4-34.6]	<0.001
Fibrinogen	214.4 ± 123.5	100.2 ± 74.2	259.1 ± 112.0	<0.001
FDP	40.0 [15.3-114.1]	83.4 [41.3-125.0]	32.2 [15.2-53.0]	<0.001

Values are expressed as the mean ± standard deviation, median [interquartile range], or number (%).

INR, international normalized ratio; FDP, fibrin degradation product.

Table 6. Multivariate analysis for predicting massive transfusion in all patients

Characteristics	Univariate OR [95% CI]	Multivariate OR [95% CI]	*P* value
Age	1.057 [1.011-1.104]		
Initial shock index	28.645 [14.347-57.193]	18.093 [5.658-57.864]	<0.001
Hemoglobin	0.750 [0.691-0.815]		
INR	20.304 [10.514-39.209]	7.103 [2.998-16.831]	<0.001
Albumin	0.085 [0.052-0.139]	0.343[0.124-0.952]	0.04
Lactate	1.782 [1.551-2.046]	1.217 [1.028-1.439]	0.02
D-dimer	1.007 [1.004-1.009]		
Fibrinogen	0.986 [0.983-0.989]	0.990 [0.986-1.002]	<0.001
FDP	1.001 [1.000-1.002]		

Multivariate analysis was conducted using a backward stepwise logistic regression model.
OR, odds ratio; CI, confidential interval; INR, international normalized ratio; FDP, fibrin degradation product.

Table 7. Multivariate analysis for predicting massive transfusion in the training set

Characteristics	Univariate OR [95% CI]	Multivariate OR [95% CI]	*P* value
Age	1.068 [1.015-1.123]		
Initial shock index	23.232 [10.619-50.825]	10.186 [2.862-36.259]	<0.001
Hemoglobin	0.765 [0.695-0.842]		
INR	40.049 [15.822-101.370]		
Albumin	0.084 [0.047-0.152]		
Lactate	1.794 [1.526-2.109]		
D-dimer	1.006 [1.003-1.009]		
Fibrinogen	0.985 [0.982-0.988]	0.990 [0.985-1.002]	<0.001
FDP	1.001 [1.000-1.002]		

Multivariate analysis was conducted using a backward stepwise logistic regression model.
OR, odds ratio; CI, confidential interval; INR, international normalized ratio; FDP, fibrin degradation product.

Table 8. Multivariate analysis for predicting massive transfusion in the validation set

Characteristics	Univariate OR [95% CI]	Multivariate OR [95% CI]	p-value
Age	1.019 [0.928-1.119]		
Initial shock index	54.183 [12.507-234.739]	135.061 [5.673-3215.326]	<0.001
Hemoglobin	0.711 [0.604-0.837]		
INR	6.894 [2.692-17.658]		
Albumin	0.086 [0.034-0.215]		
Lactate	1.754 [1.342-2.294]		
D-dimer	1.008 [1.003-1.014]		
Fibrinogen	0.988 [0.983-0.993]		
FDP	1.002 [1.000-1.004]		

Multivariate analysis was conducted using a backward stepwise logistic regression model.
OR, odds ratio; CI, confidential interval; INR, international normalized ratio; FDP, fibrin degradation product.

Table 9. Machine learning analysis for predicting massive transfusion with early features in the training set and validation set

Model	**Training set** / validation set	**AUC**	**Accuracy**	**Sensitivity**	**Specificity**	**PPV**	**NPV**	**F1**
SVM radial		0.838	0.806	0.692	0.850	0.642	0.876	0.666
		0.783	0.746	0.651	0.783	0.538	0.852	0.589
SVM poly		0.782	0.784	0.538	0.880	0.636	0.830	0.583

	0.751	0.694	0.651	0.711	0.466	0.840	0.543
KNN	0.696	0.666	0.423	0.761	0.407	0.772	0.415
	0.706	0.688	0.627	0.711	0.457	0.831	0.529
Logistic	0.816	0.795	0.730	0.820	0.612	0.887	0.666
	0.782	0.720	0.697	0.729	0.500	0.861	0.582
XGBoost	0.799	0.752	0.692	0.776	0.545	0.866	0.610
	0.763	0.745	0.651	0.783	0.538	0.852	0.589
Random Forest	0.784	0.741	0.692	0.761	0.529	0.864	0.600
	0.755	0.740	0.534	0.819	0.500	0.861	0.534
Decision tree	0.722	0.709	0.692	0.716	0.486	0.857	0.571
	0.729	0.714	0.720	0.711	0.492	0.868	0.584

Table 10. Machine learning analysis for predicting massive transfusion with all features in the training set and validation set

Model	**Training set**	**AUC**	**Accuracy**	**Sensitivity**	**Specificity**	**PPV**	**NPV**	**F1**

validation set							
SVM radial	0.898	0.870	0.807	0.895	0.750	0.923	0.777
	0.884	0.876	0.790	0.909	0.772	0.918	0.781
SVM poly	0.897	0.881	0.807	0.910	0.777	0.924	0.792
	0.837	0.826	0.651	0.909	0.736	0.870	0.691
KNN	0.860	0.870	0.730	0.925	0.791	0.898	0.760
	0.796	0.831	0.627	0.909	0.729	0.863	0.675
Logistic	0.924	0.892	0.846	0.910	0.785	0.938	0.924
	0.884	0.883	0.767	0.927	0.804	0.911	0.785
XGBoost	0.906	0.903	0.807	0.940	0.840	0.926	0.823
	0.883	0.870	0.697	0.936	0.810	0.888	0.750
Random Forest	0.901	0.892	0.846	0.910	0.785	0.938	0.814
	0.885	0.896	0.744	0.954	0.864	0.905	0.800
Decision tree	0.874	0.849	0.807	0.865	0.700	0.920	0.750
	0.793	0.805	0.767	0.819	0.622	0.900	0.687

Figure 1. Flow diagram illustrating the study selection process

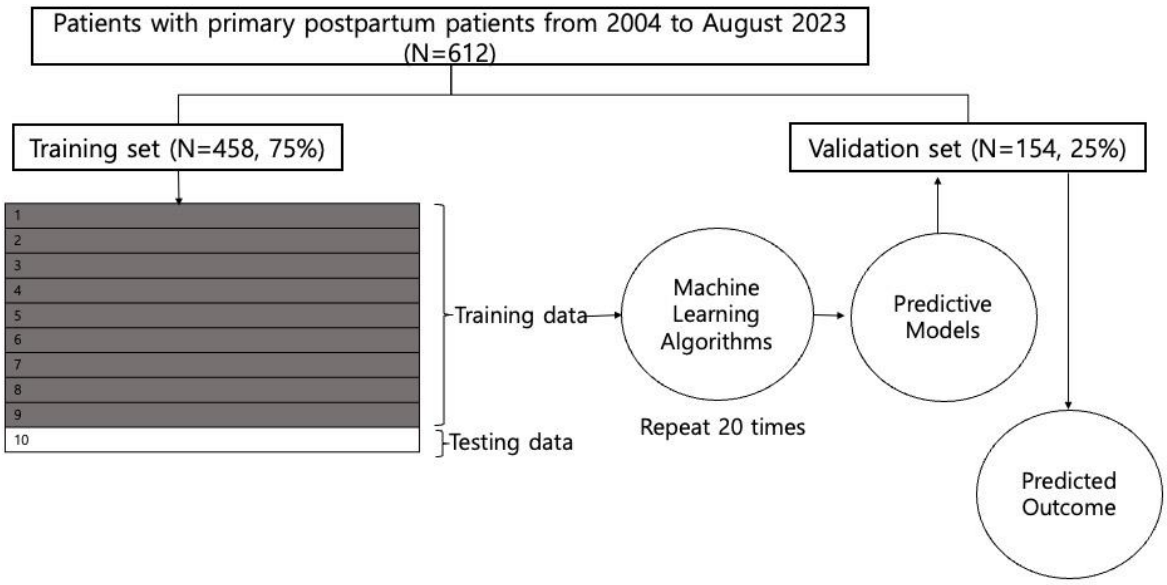
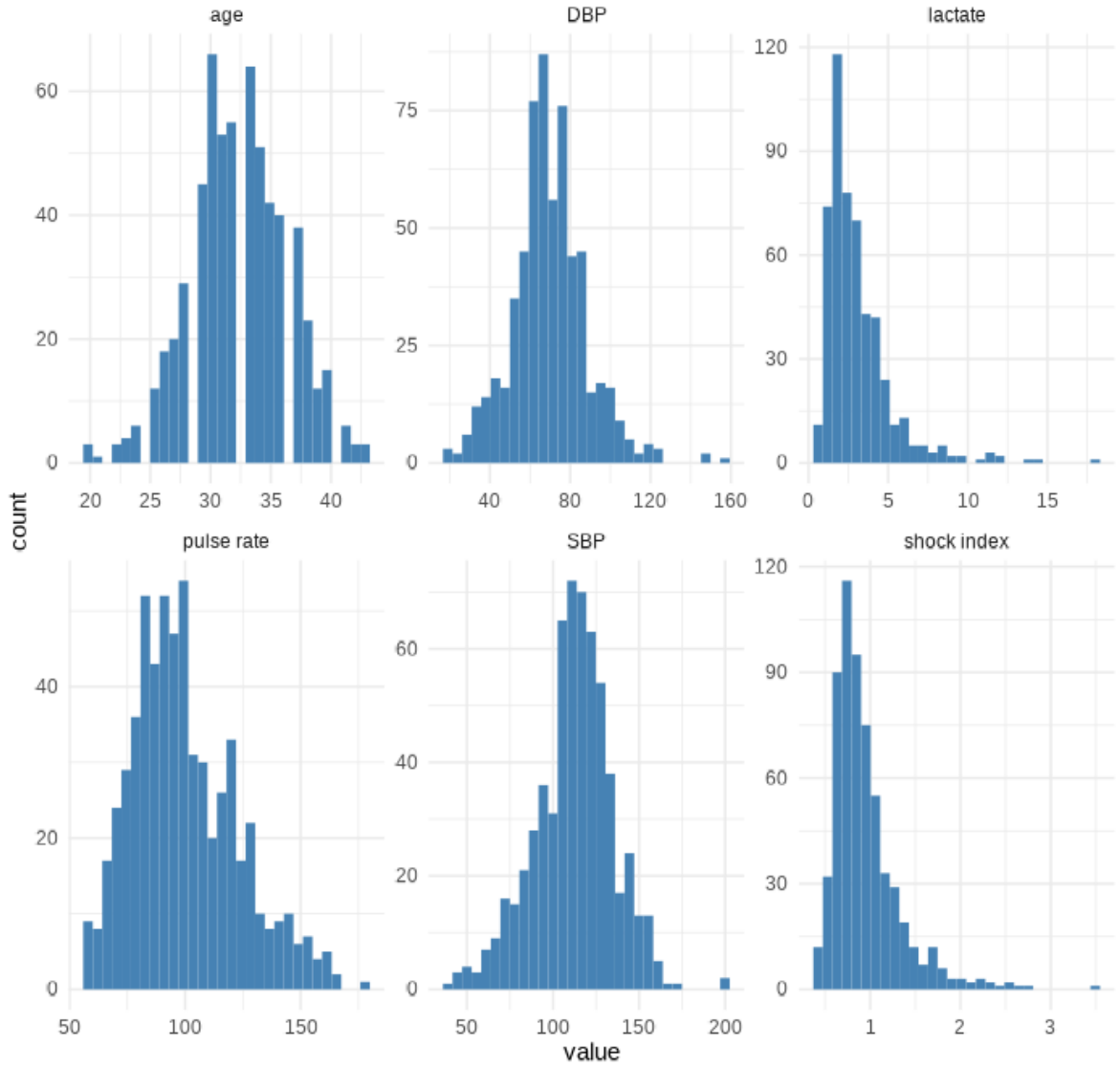


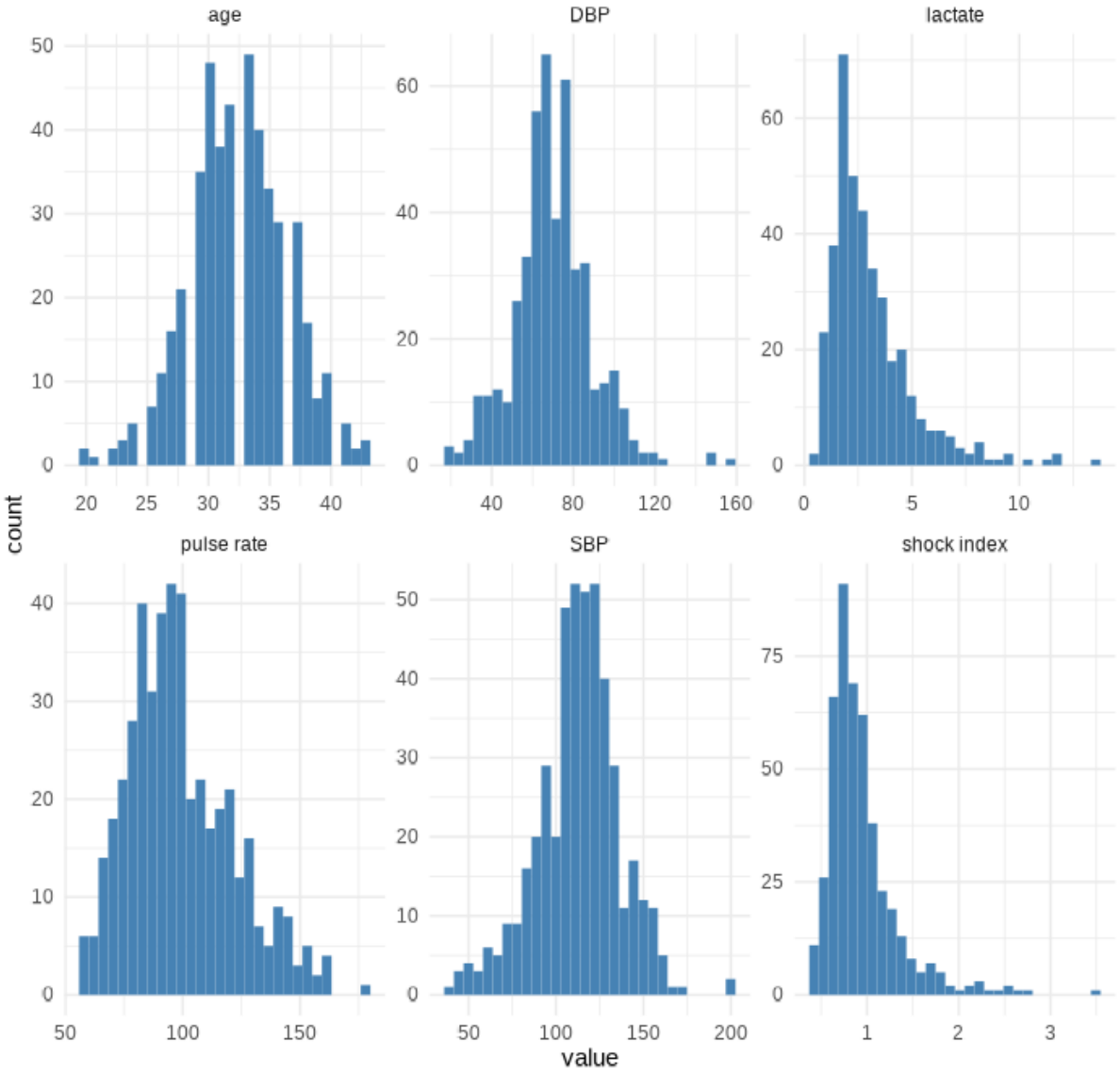
Figure 2. Distributions of numerical covariates in the training set (a), validation set (b), and all study patients (c) for early features (A) and all features (B)

(A-a) Early features in all study patients



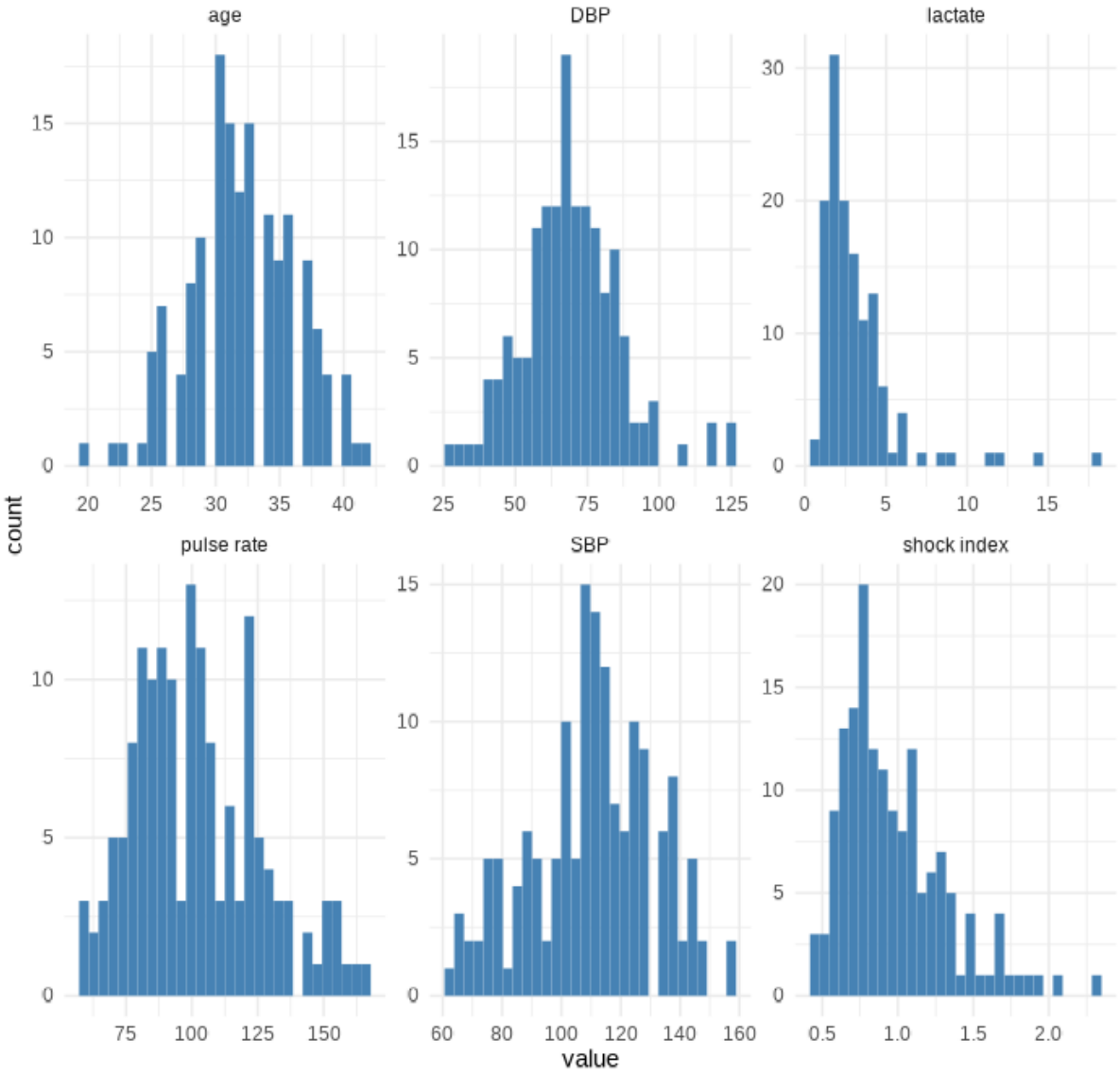
DBP, diastolic blood pressure; SBP, systolic blood pressure.

(A-b) Early features in the training set

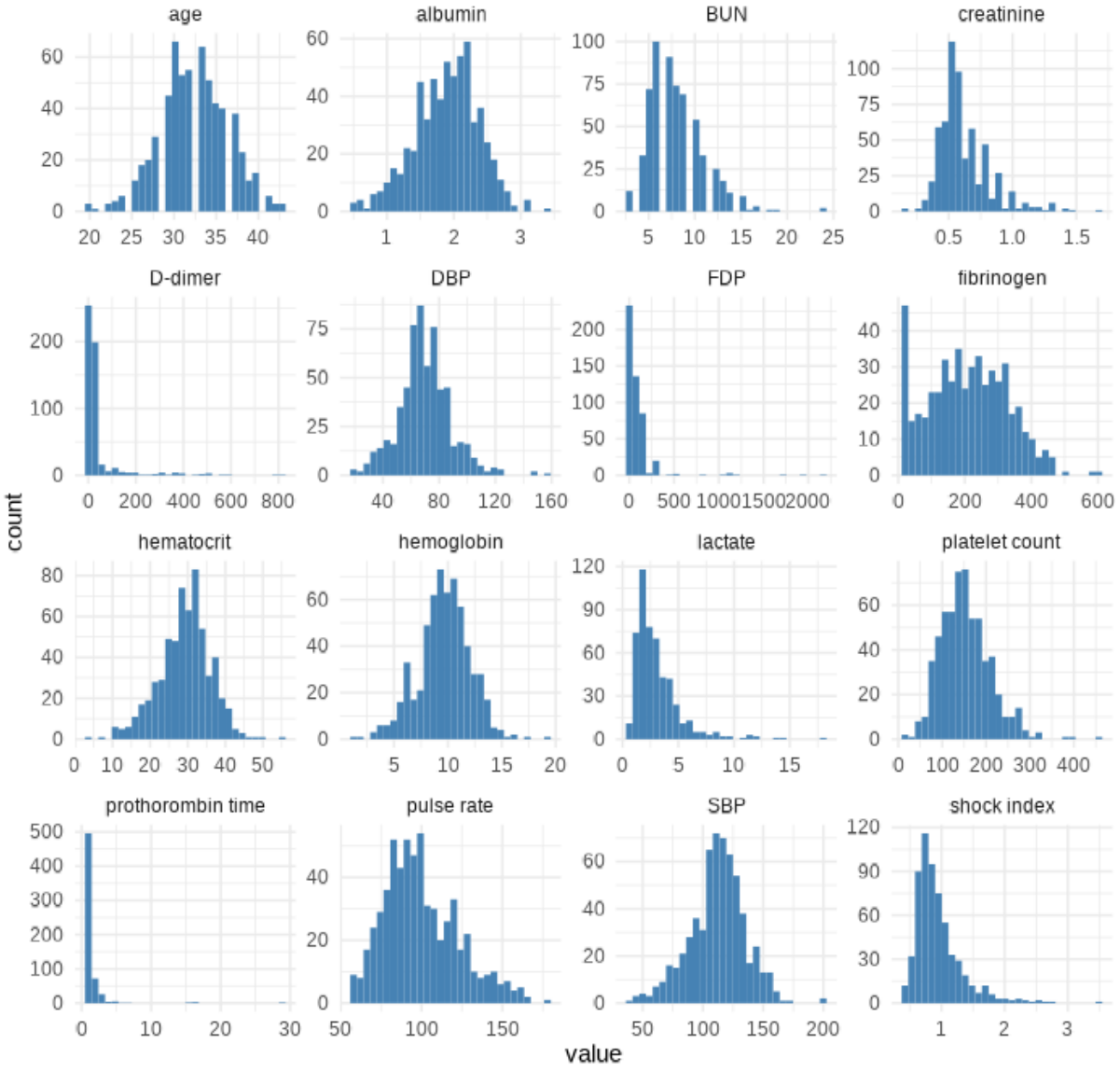


DBP, diastolic blood pressure; SBP, systolic blood pressure.

(A-c) Early features in the validation set

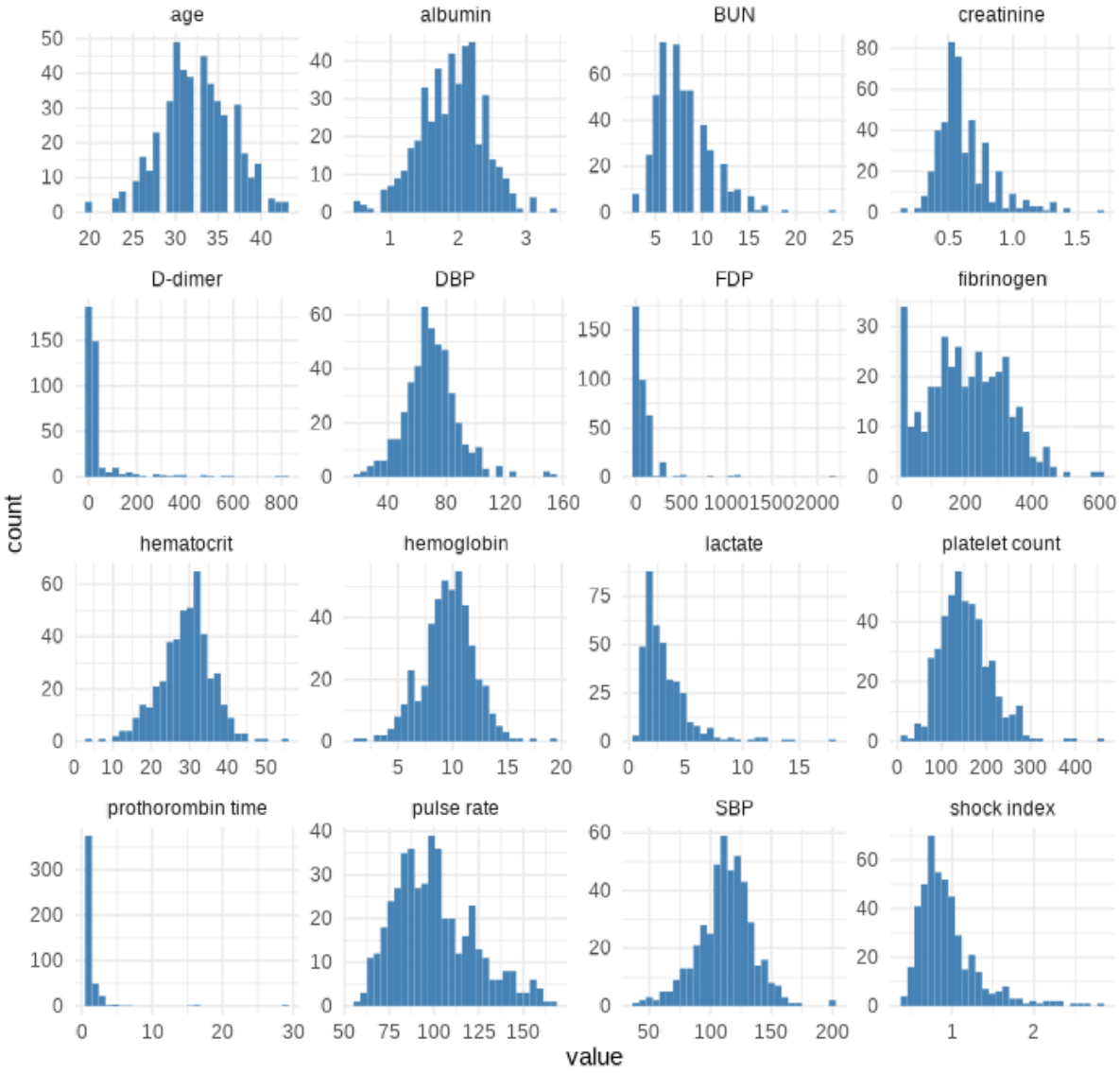


(B-a) All features in all study patients



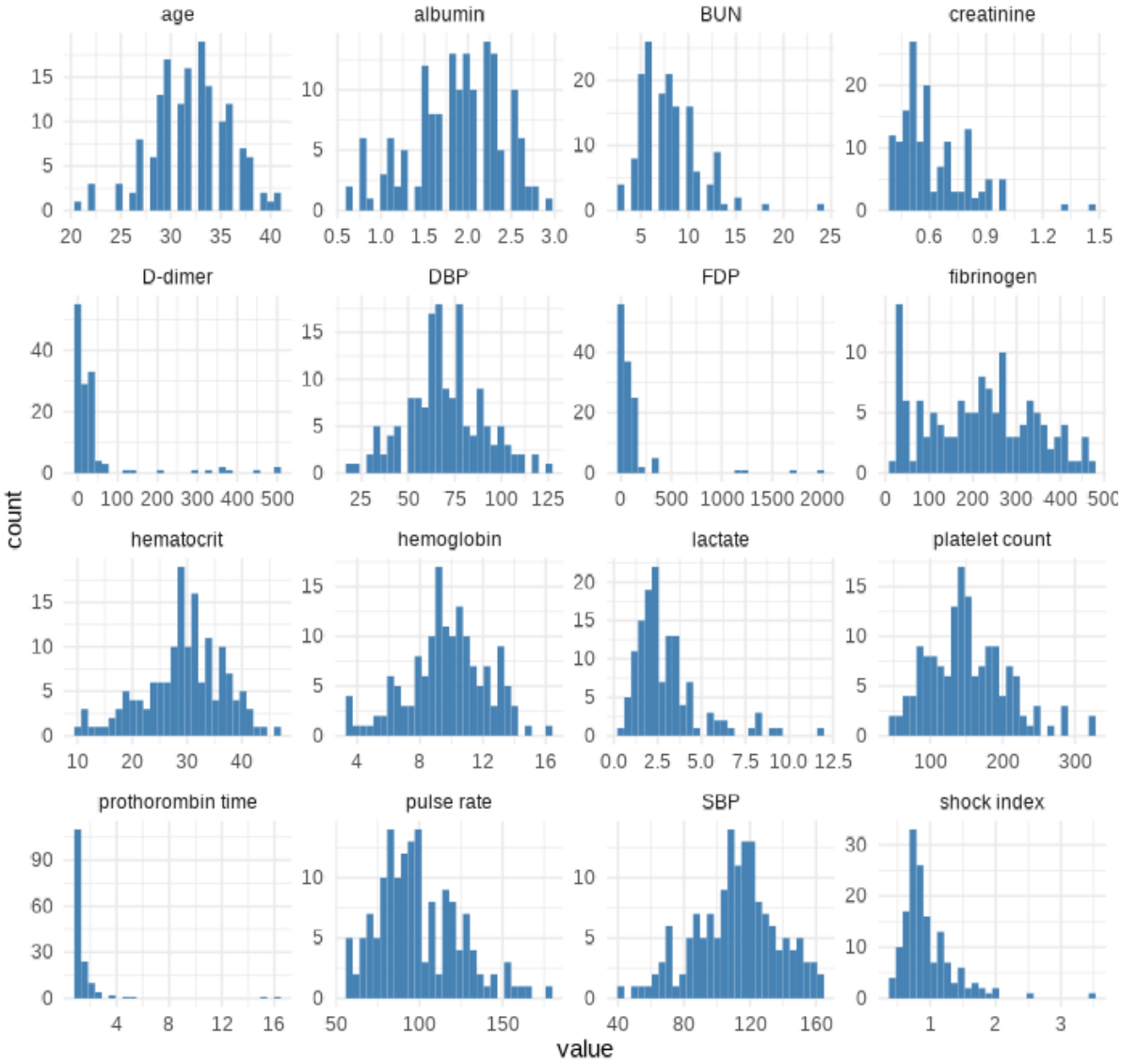
BUN, blood urea nitrogen; DBP, diastolic blood pressure; FDP, fibrin degradation product; SBP, systolic blood pressure.

(B-b) All features in the training set



BUN, blood urea nitrogen; DBP, diastolic blood pressure; FDP, fibrin degradation product; SBP, systolic blood pressure.

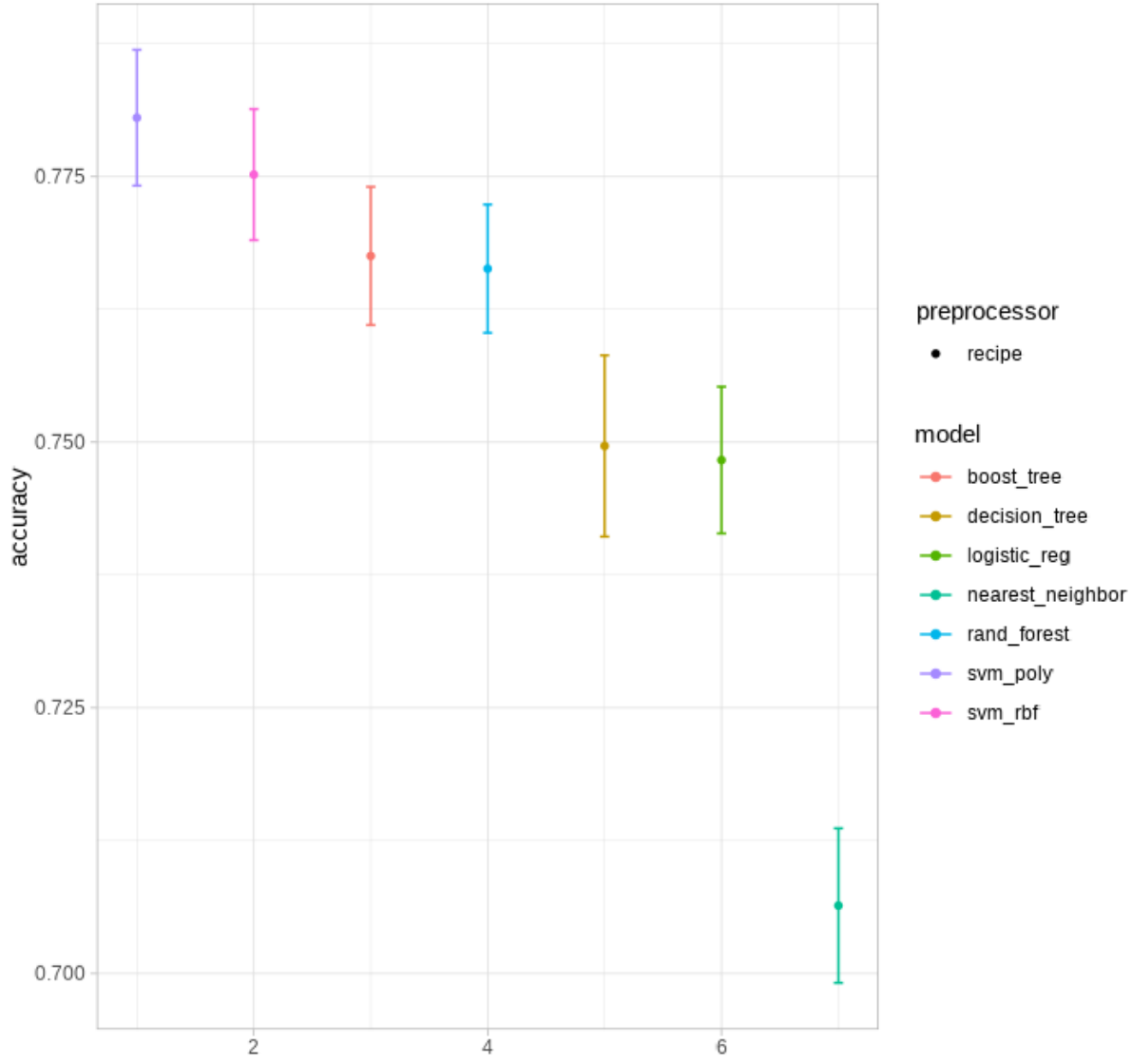
(B-c) All features in the validation set



BUN, blood urea nitrogen; DBP, diastolic blood pressure; FDP, fibrin degradation product; SBP, systolic blood pressure.

Figure 3. Estimated accuracy for the classification model tuning result using early features and all features

(A) With early features



(B) With all features

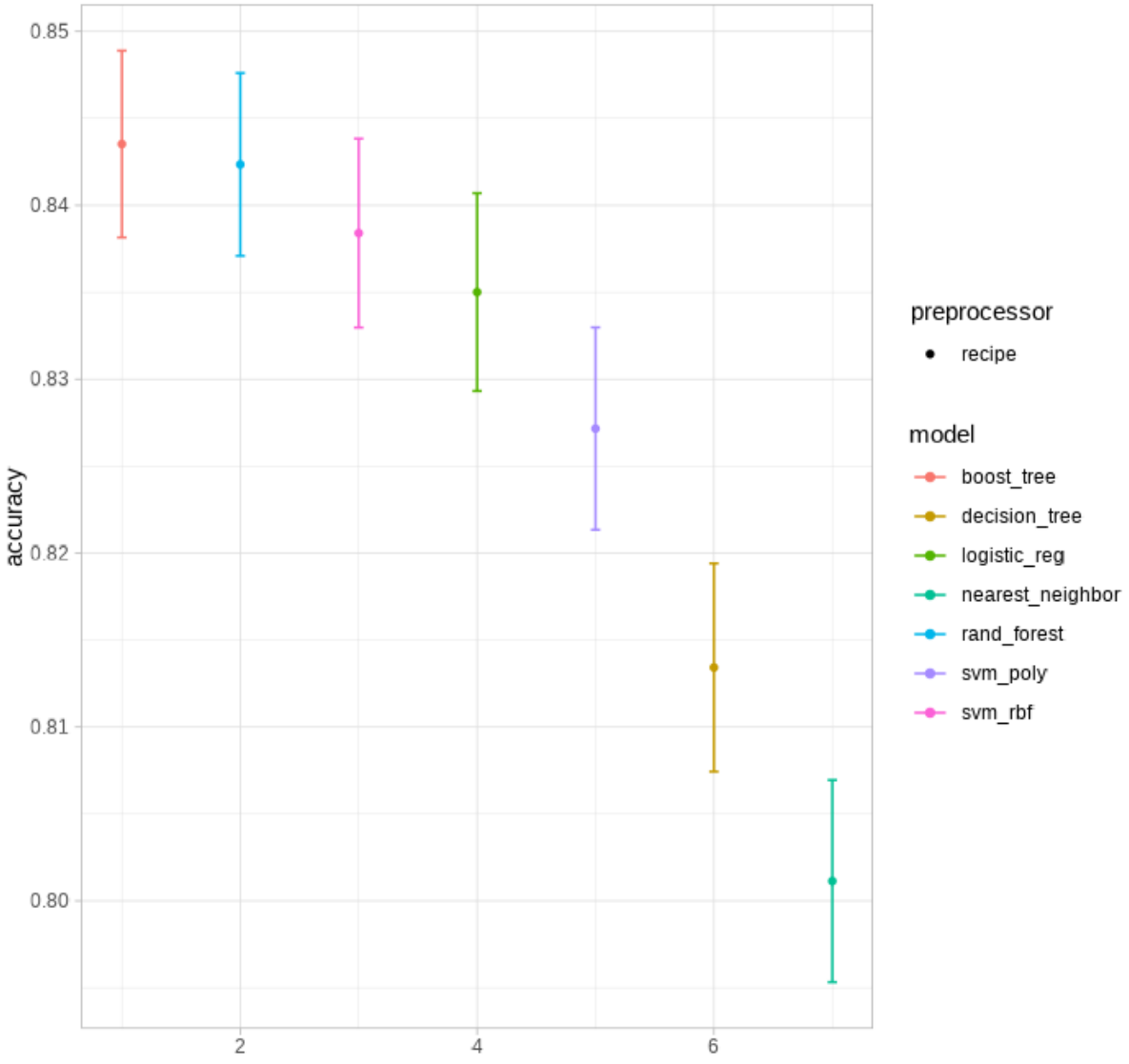
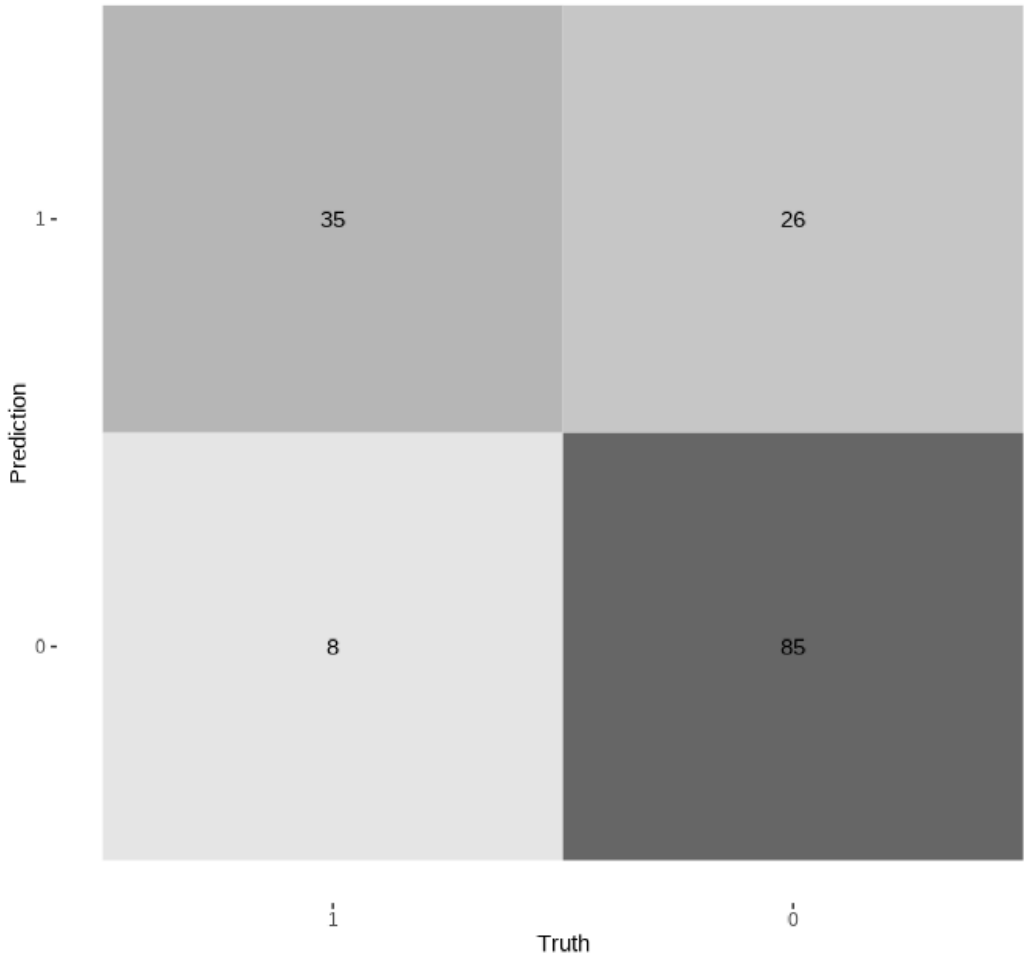
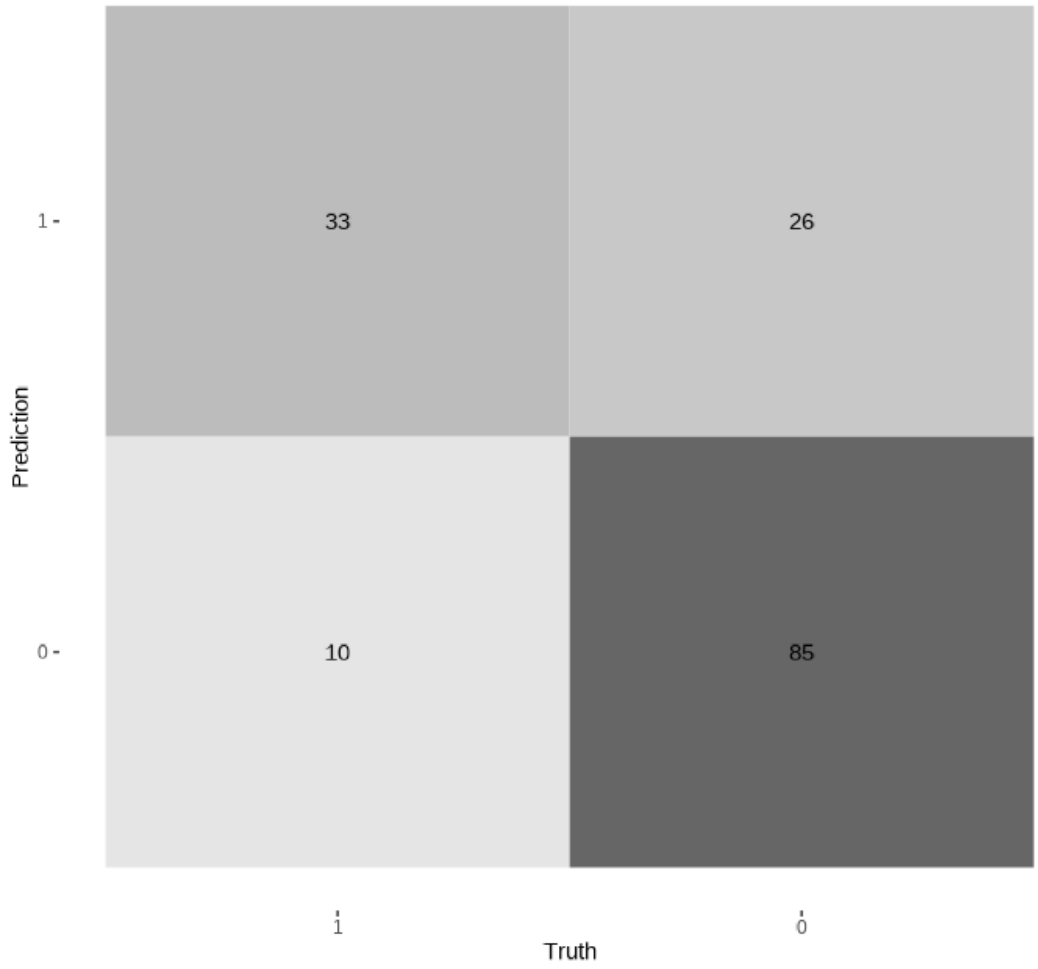


Figure 4. Confusion matrix of the prediction model with early features in each machine learning algorithm

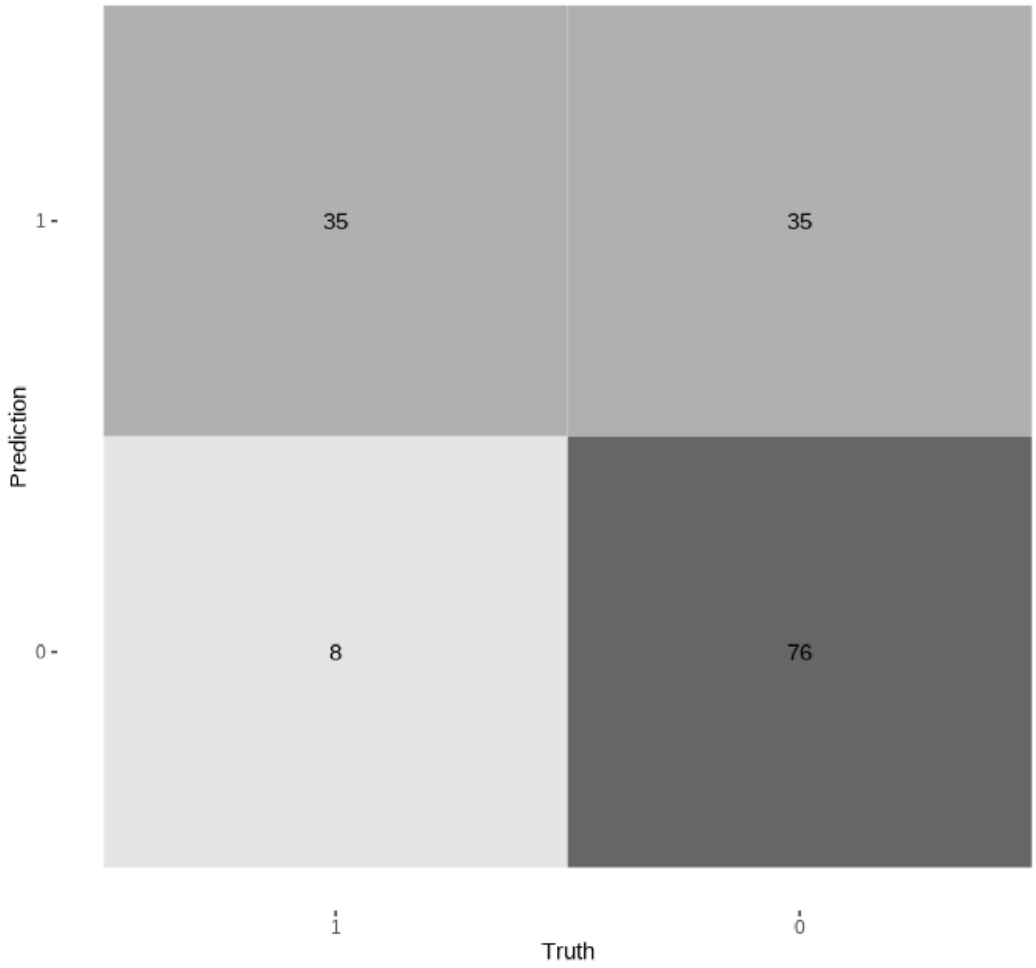
(A) XGBoost



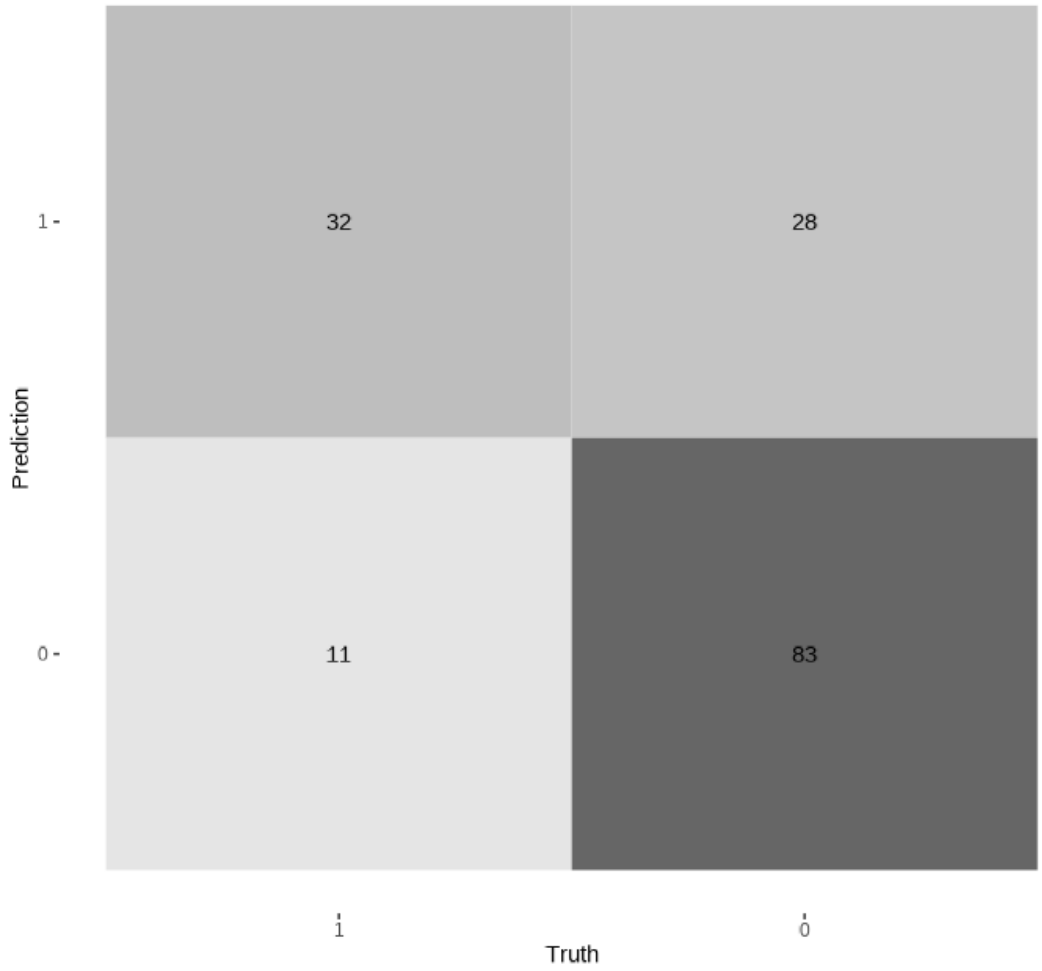
(B) Random forest



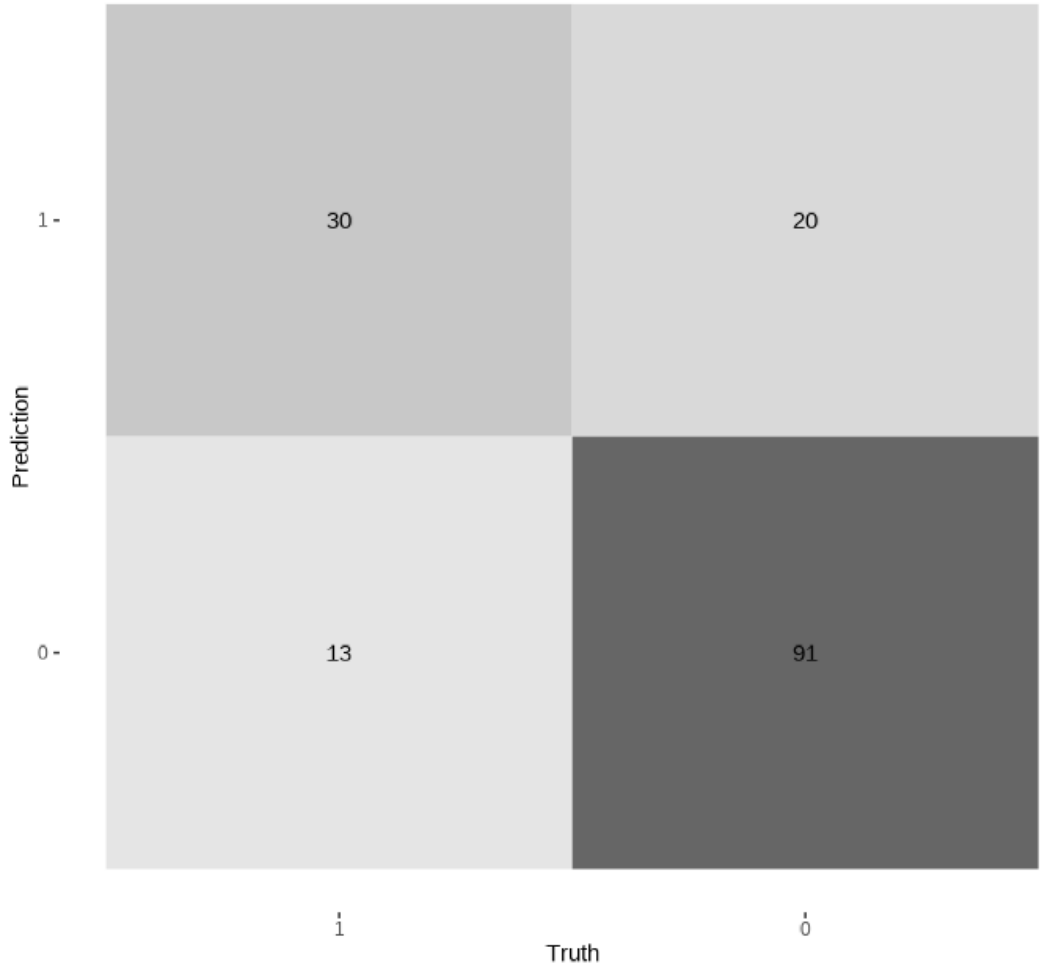
(C) Logistic Regression



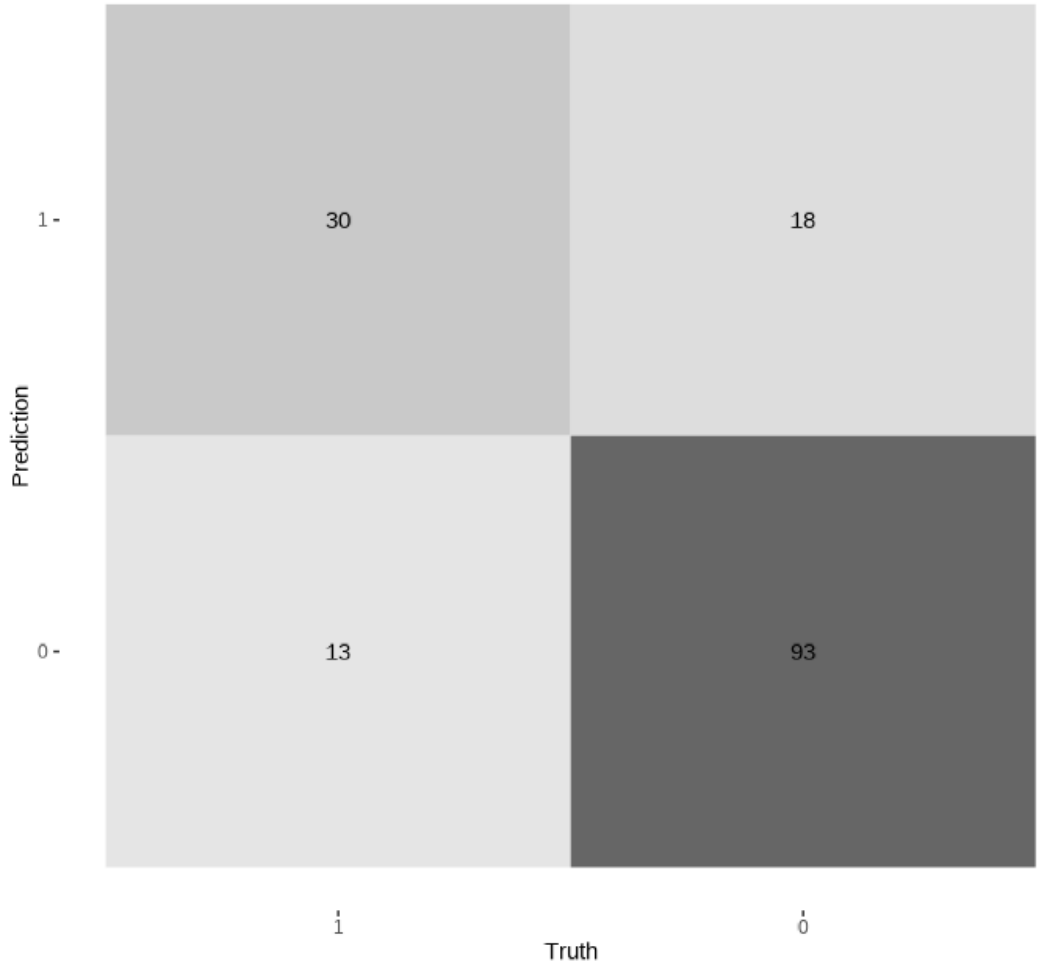
(D) KNN



(E) SVM radial



(F) SVM poly



(G) Decision tress

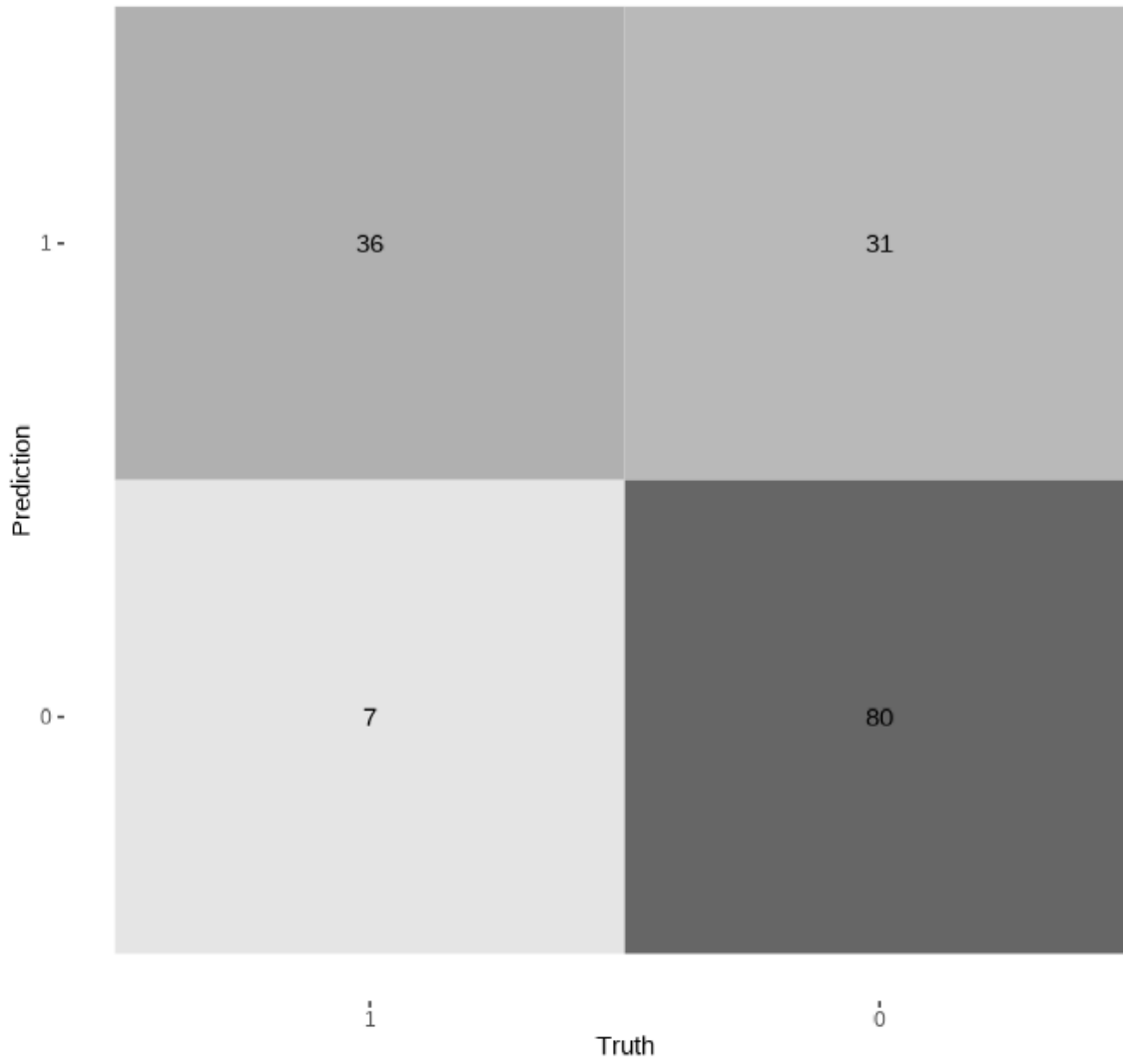
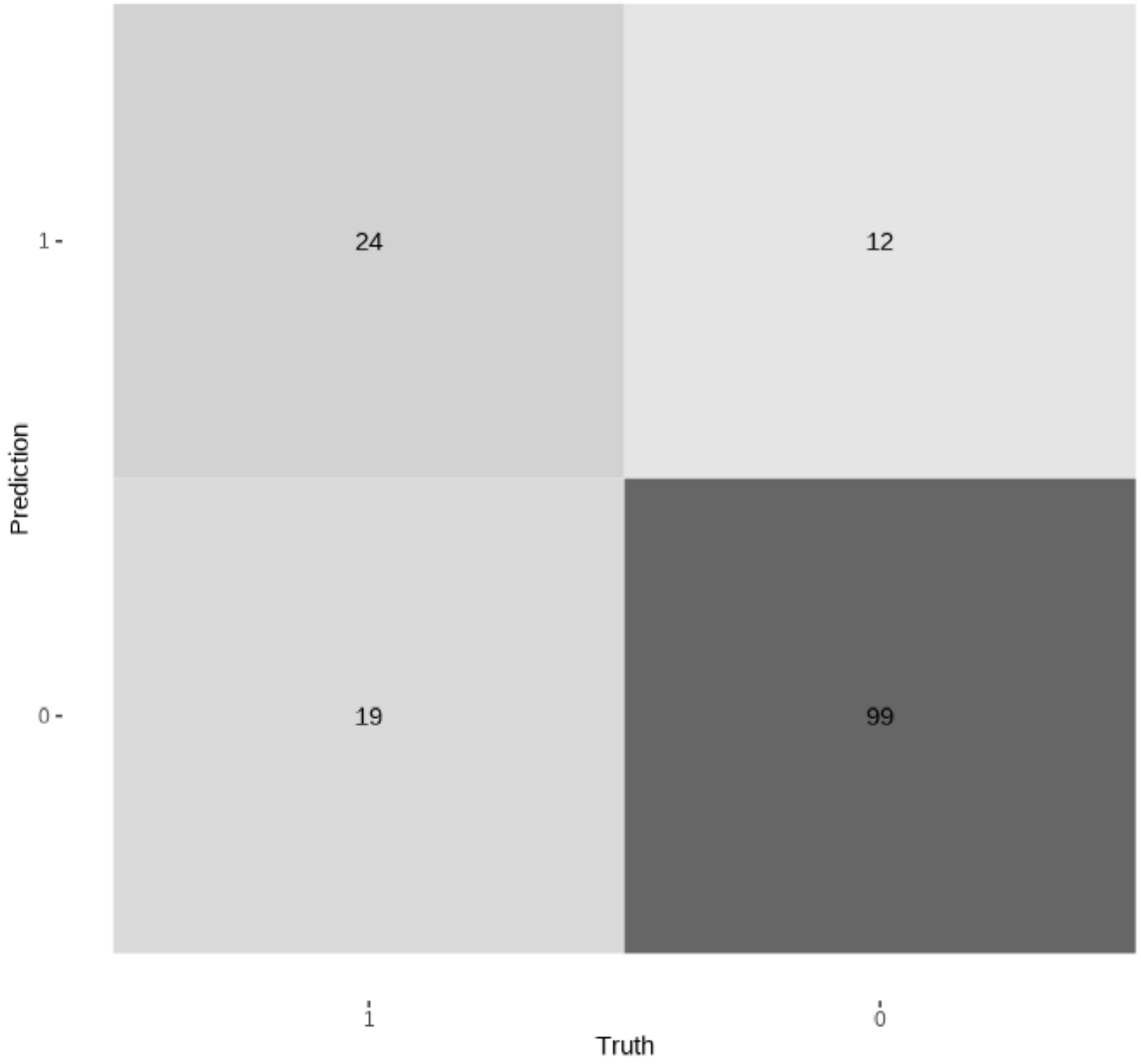
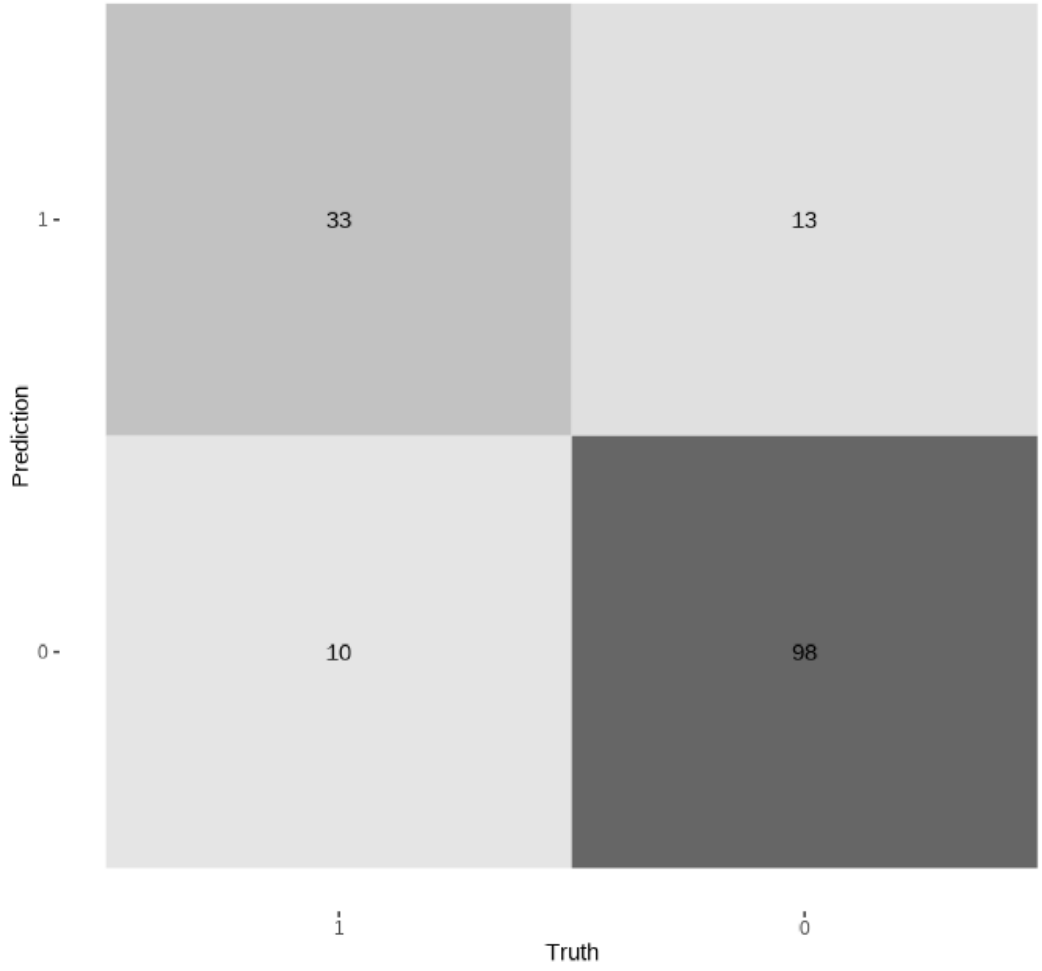


Figure 5. Confusion matrix of the prediction model with all features in each machine learning algorithm

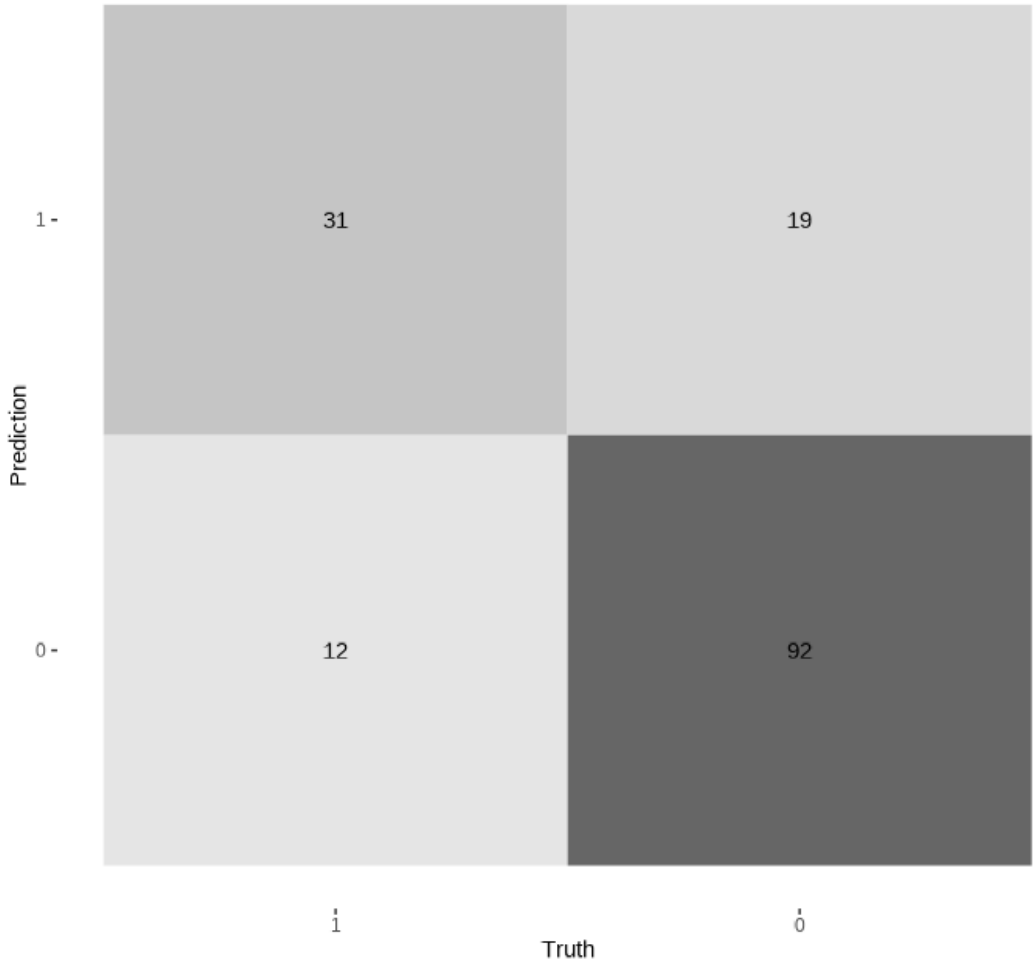
(A) XGBoost



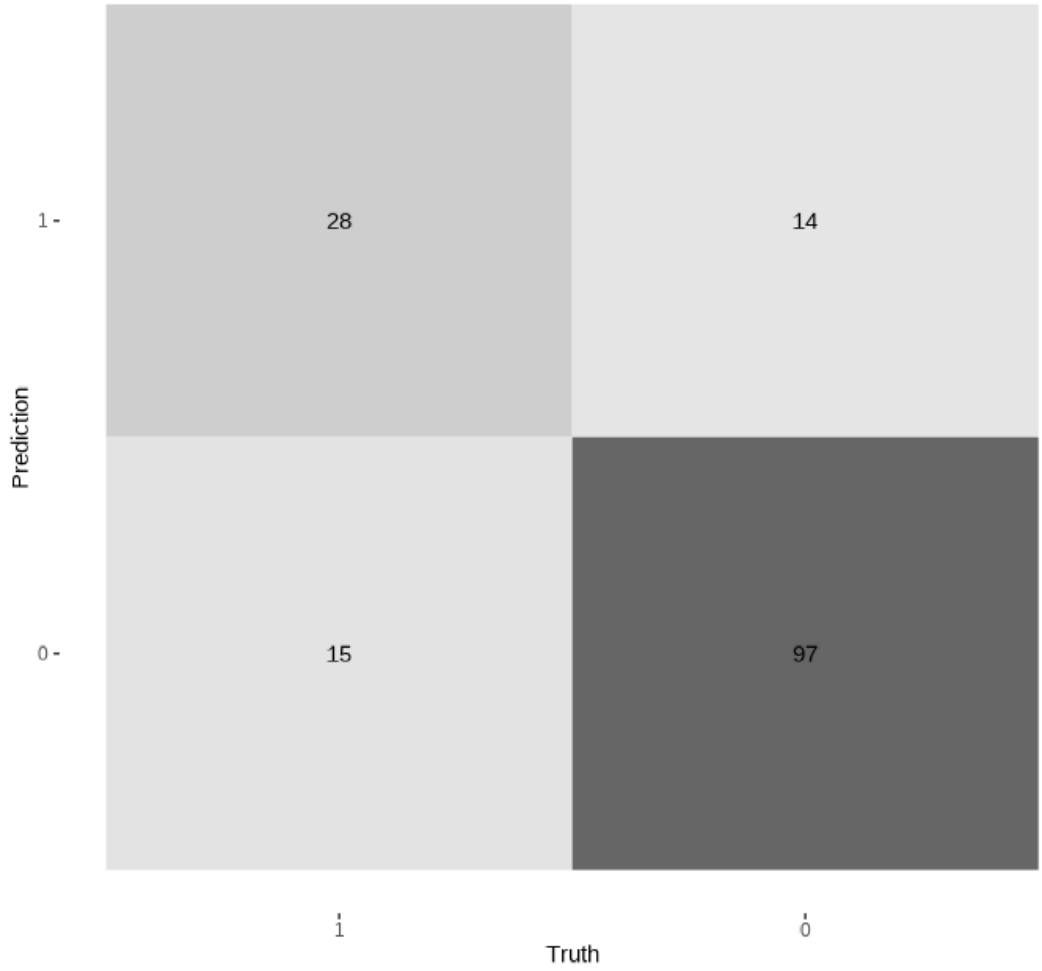
(B) Random forest



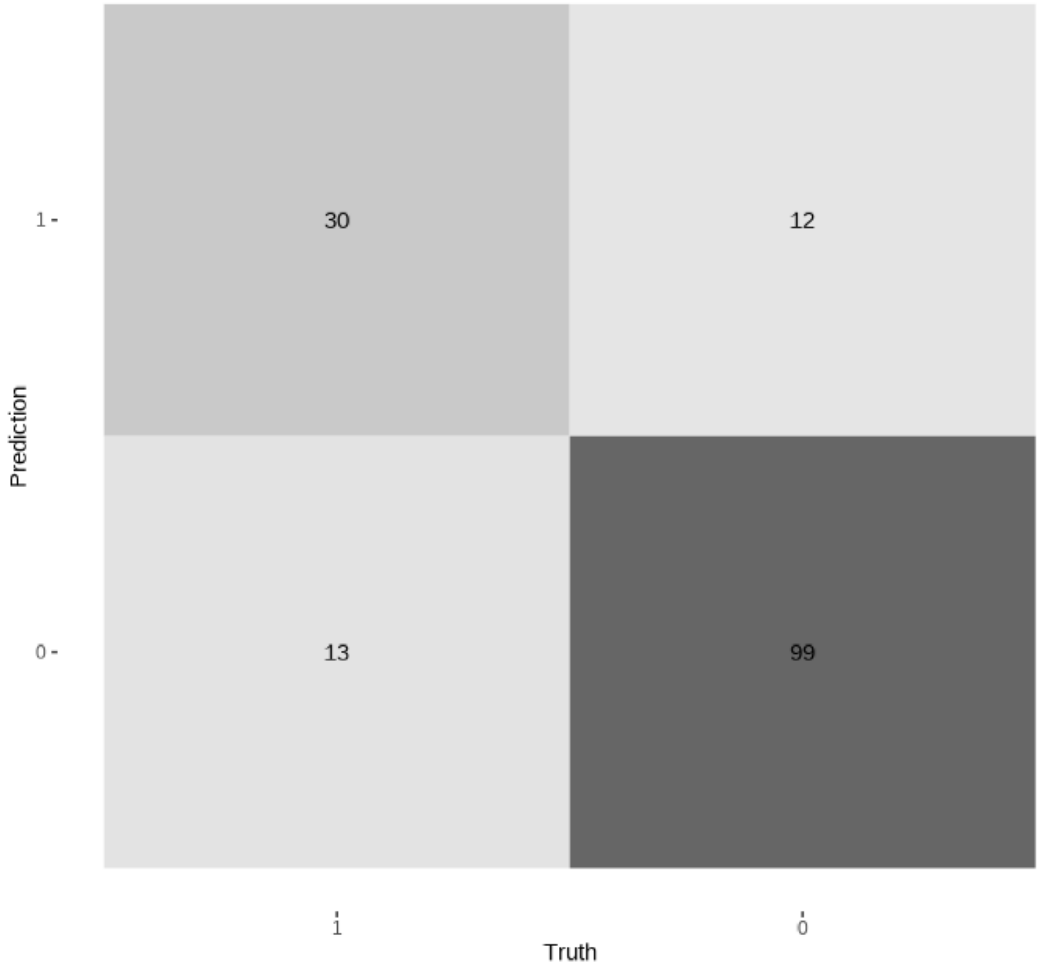
(C) Logiestic Regression



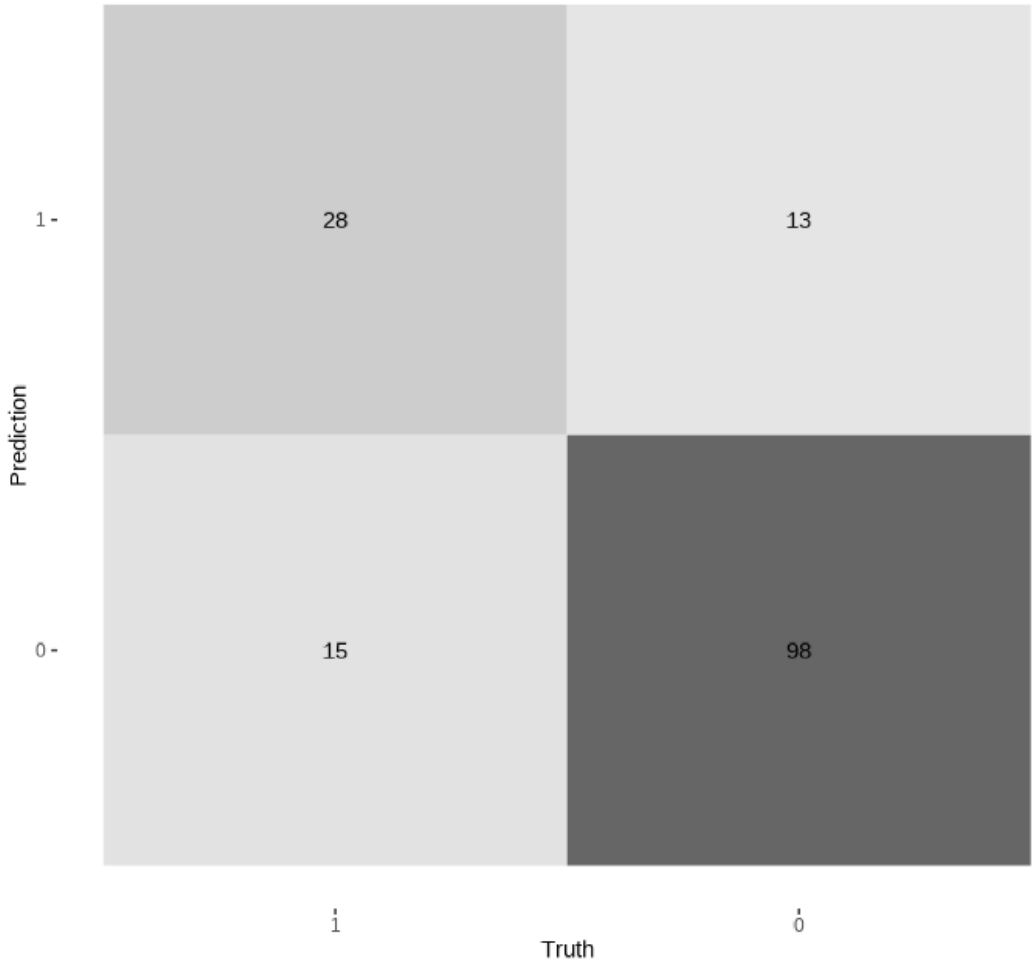
(D) KNN



(E) SVM radial



(F) SVM poly



(G) Decision tress

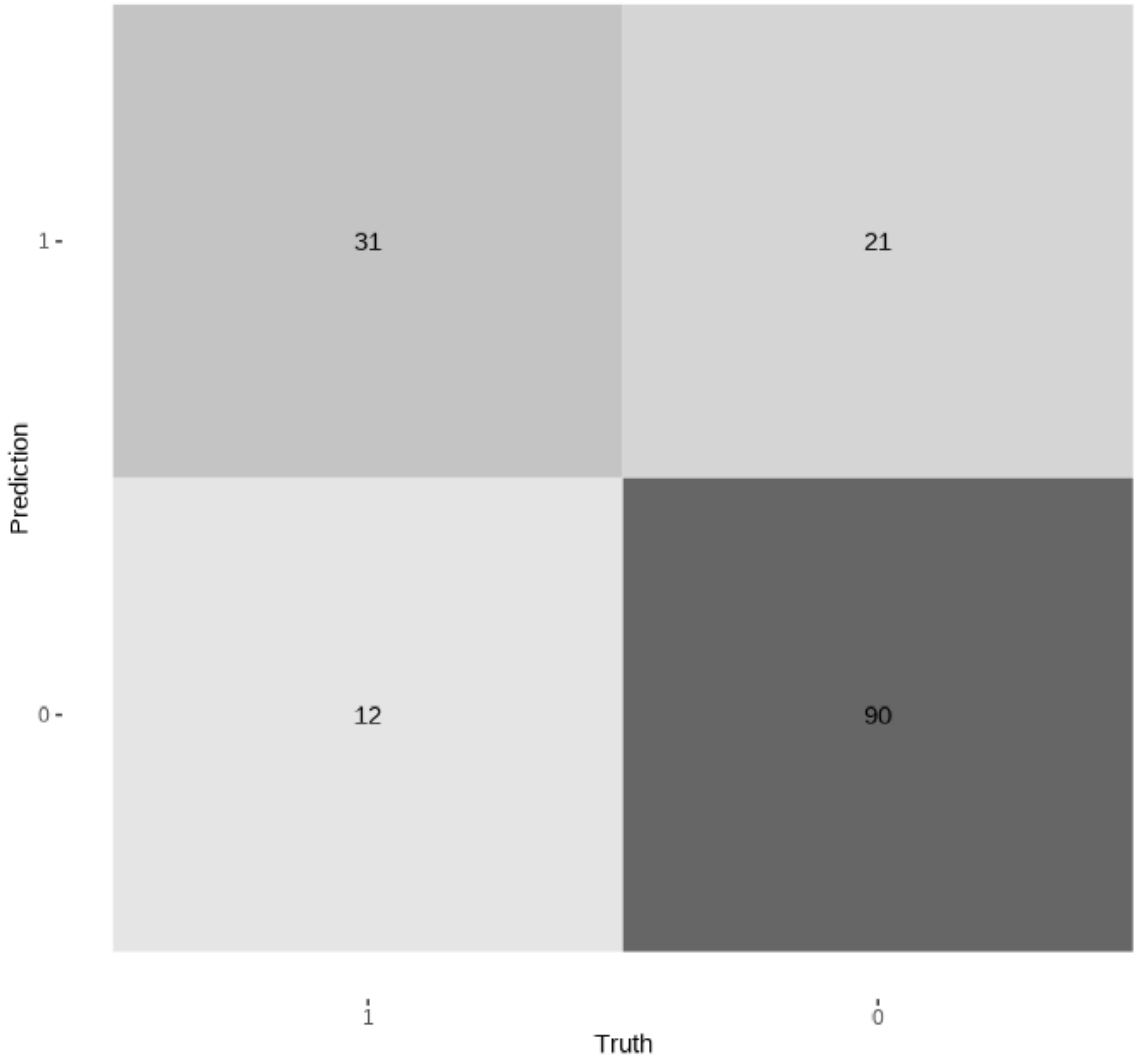
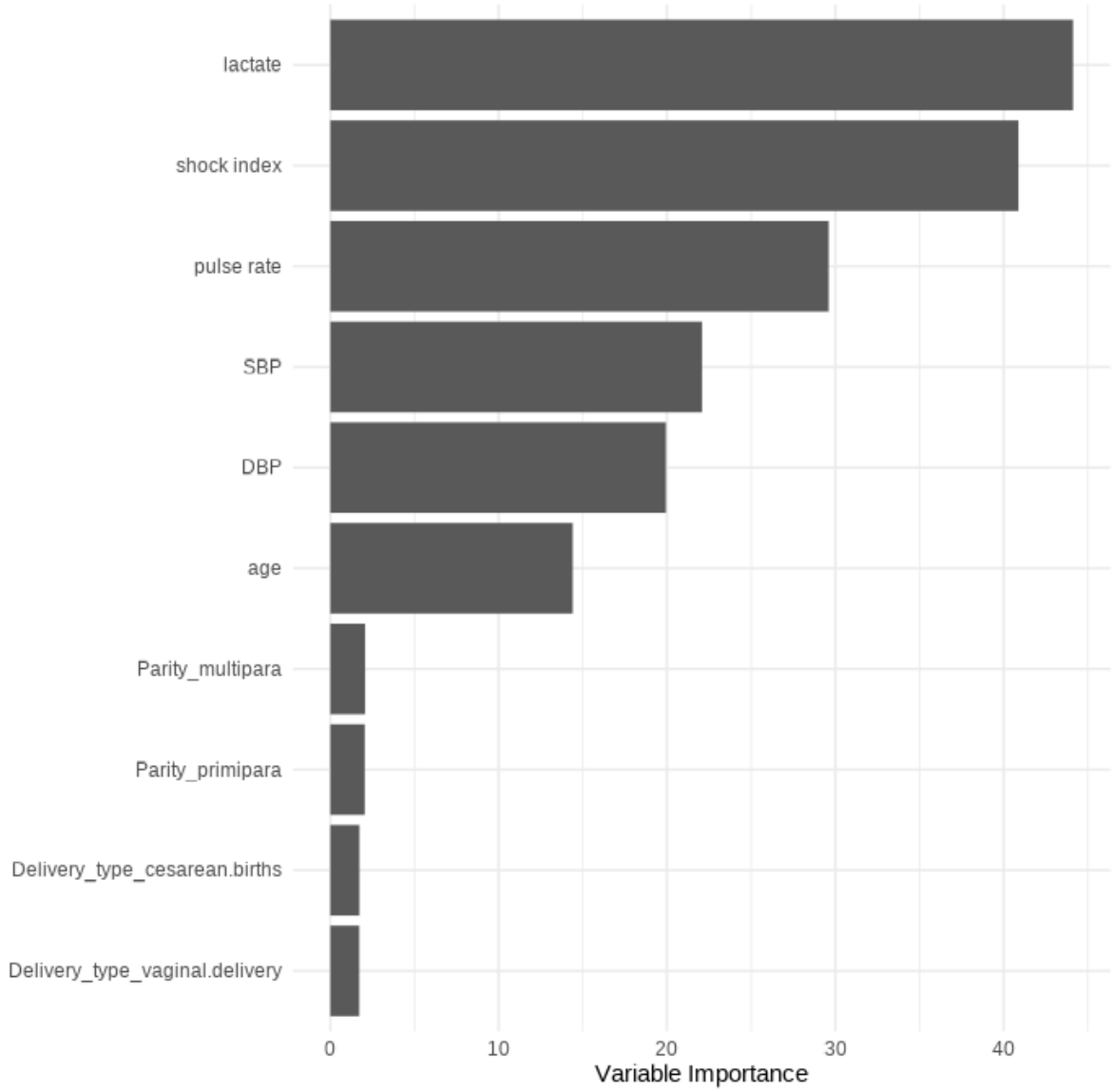


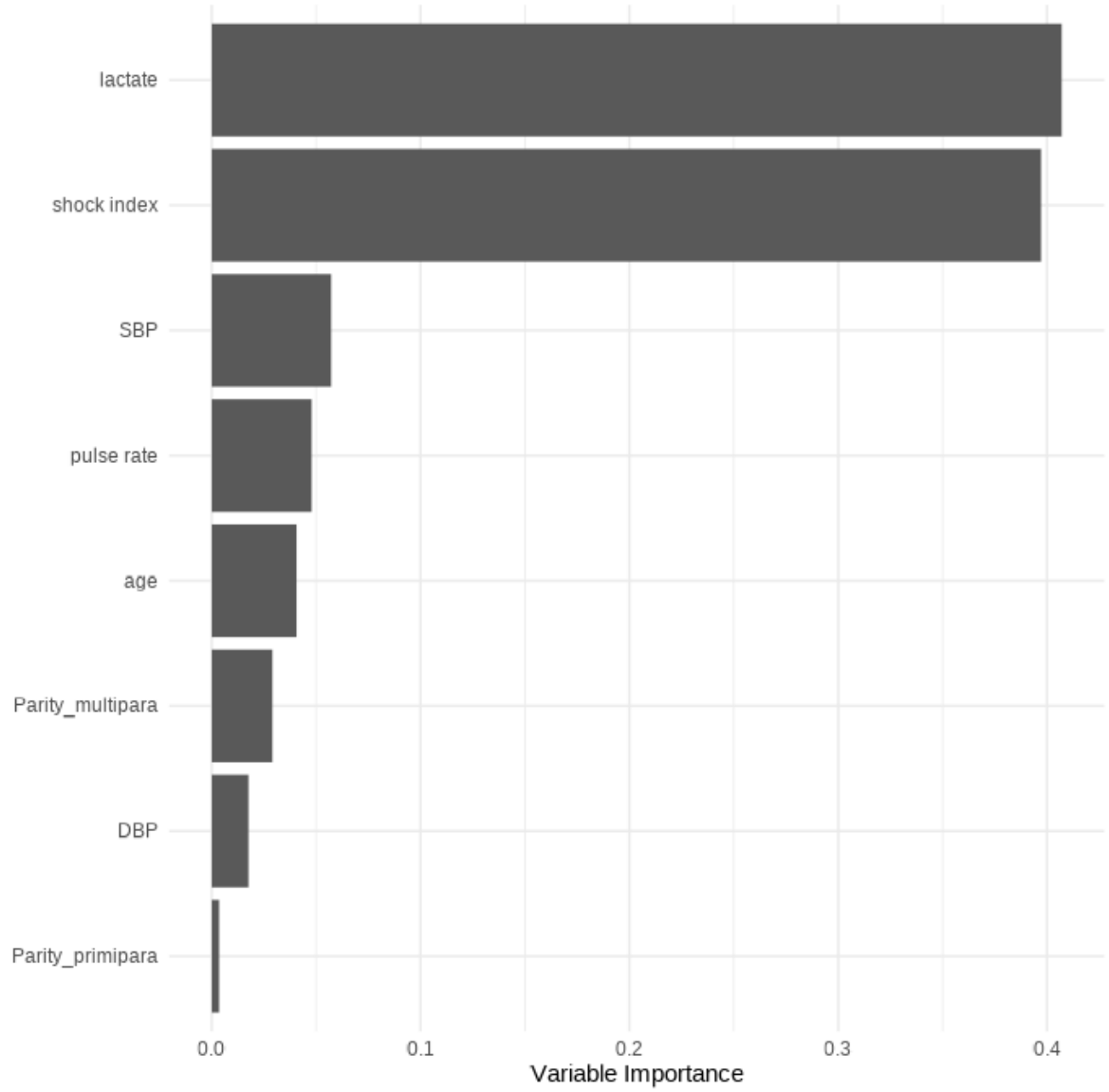
Figure 6. Plot showing the relative importance of variables for predicting massive transfusion in machine learning models of Random Forest (A), XGBoost (B) and Logistic Regression (C) using early features

(A) Random Forest



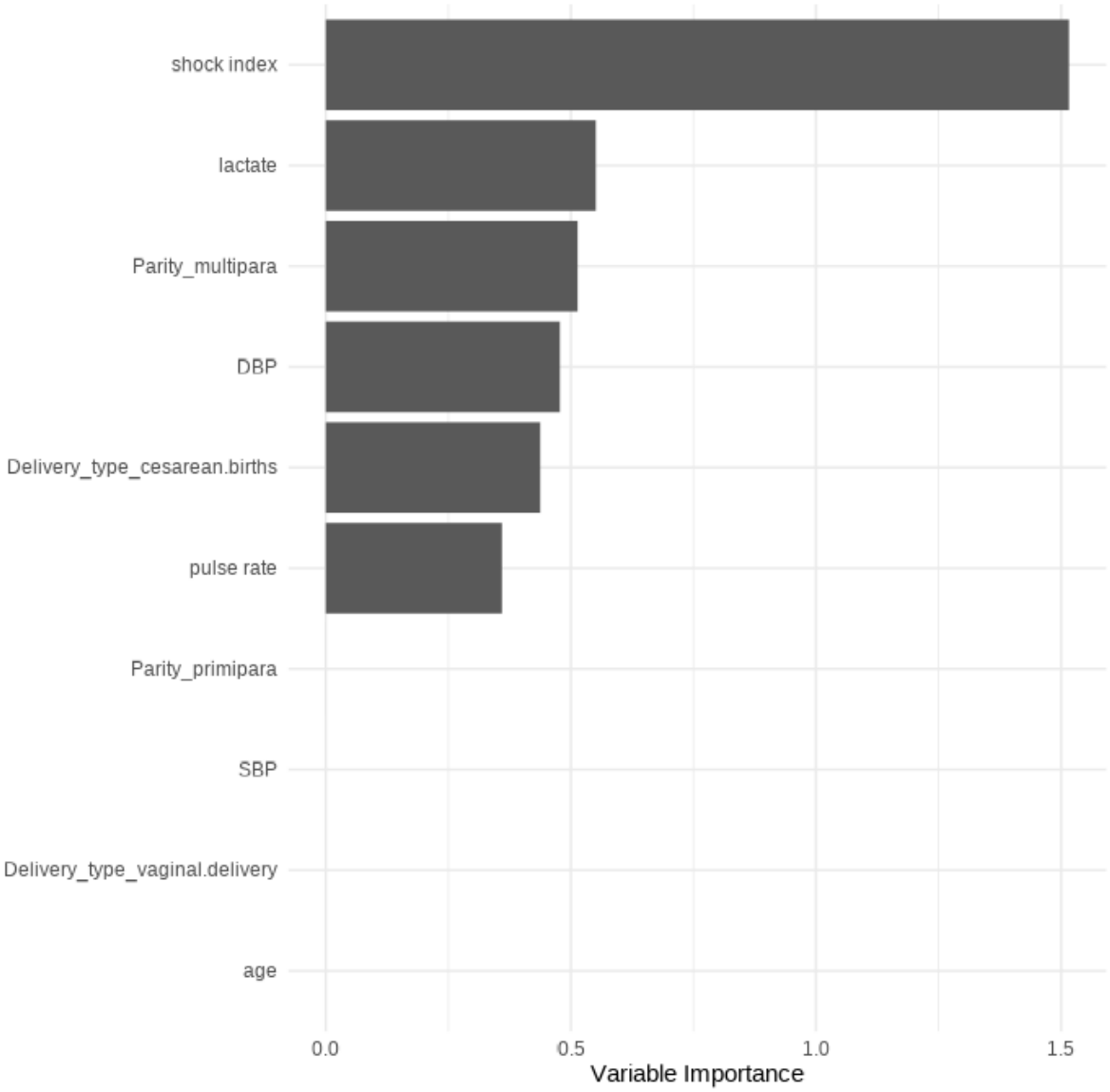
SBP, systolic blood pressure; DBP, diastolic blood pressure.

(B) XGBoost



SBP, systolic blood pressure; DBP, diastolic blood pressure.

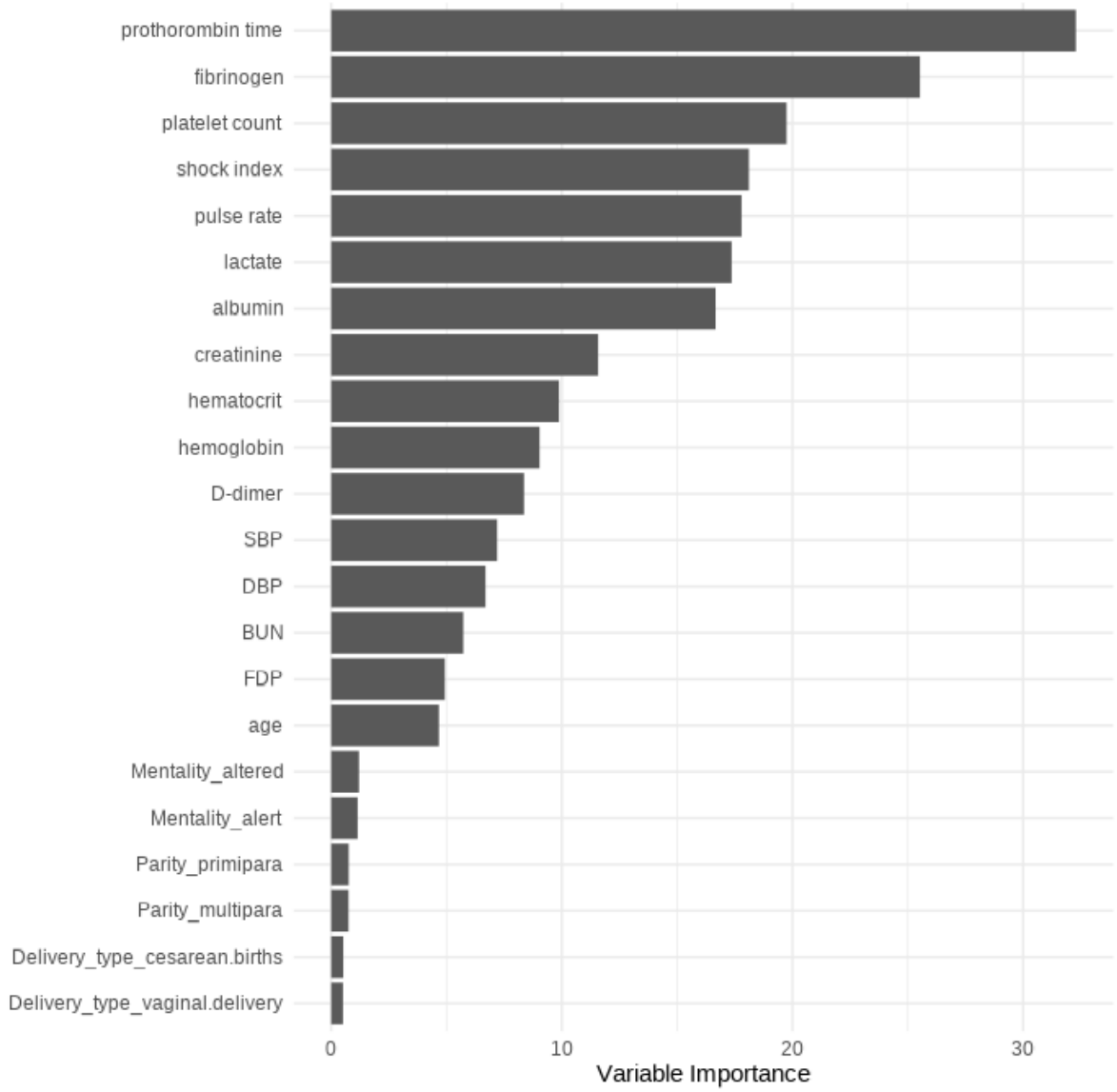
(C) Logistic Regression



DBP, diastolic blood pressure; SBP, systolic blood pressure.

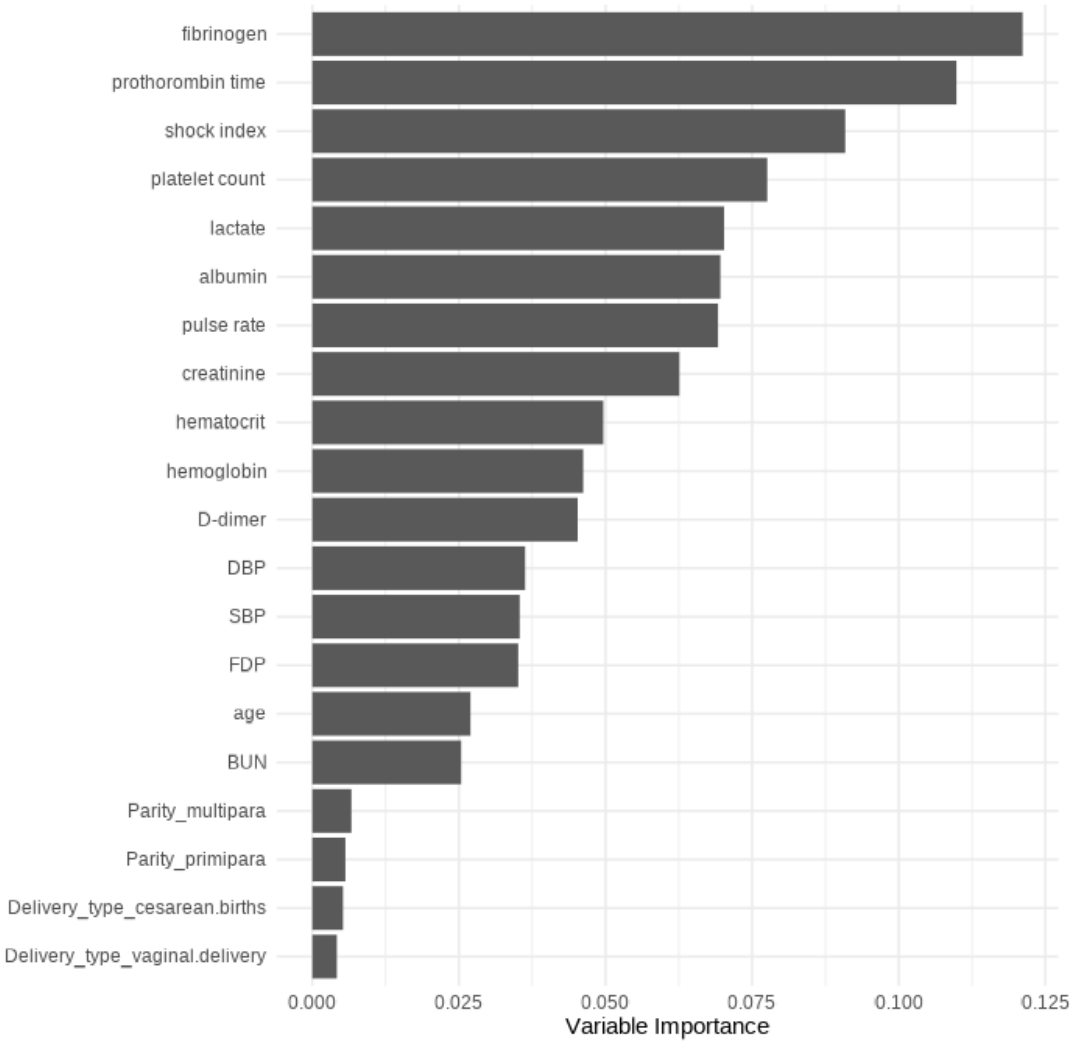
Figure 7. Plot showing the relative importance of variables for predicting massive transfusion in machine learning models of Random Forest (A), XGBoost (B) and Logistic Regression (C) using all features

(A) Random Forest



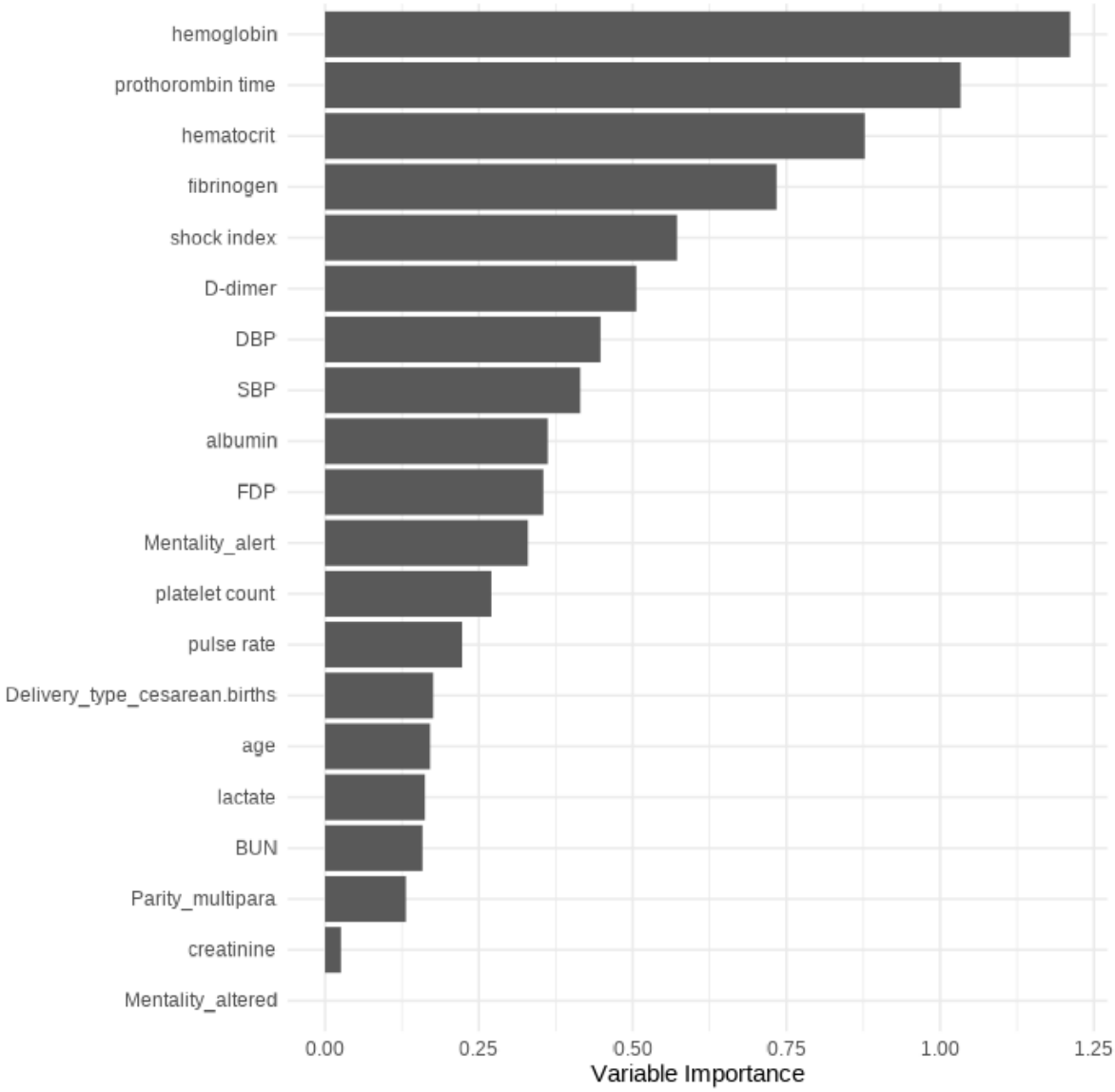
SBP, systolic blood pressure; DBP, diastolic blood pressure; BUN, blood urea nitrogen; FDP, fibrin degradation product.

(B) XGBoost



DBP, diastolic blood pressure; SBP, systolic blood pressure; FDP, fibrin degradation product; BUN, blood urea nitrogen.

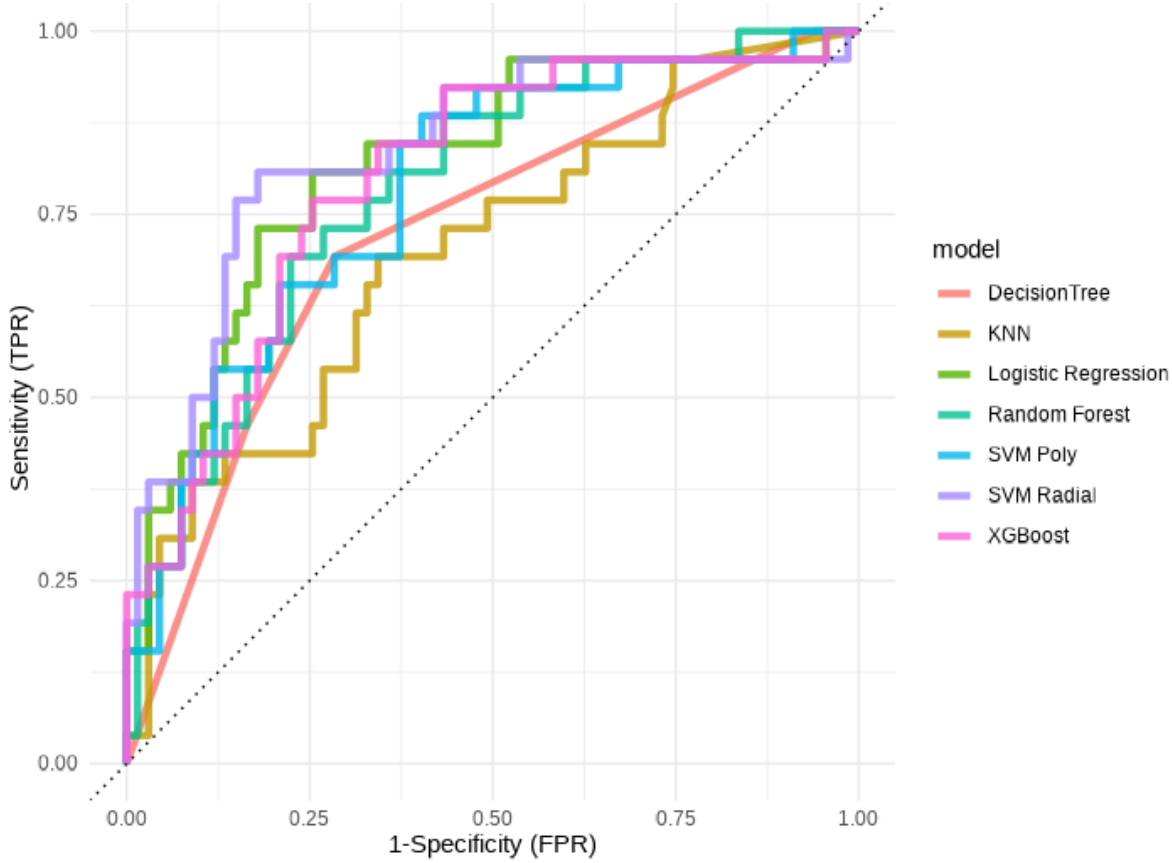
(C) Logistic Regression



DBP, diastolic blood pressure; SBP, systolic blood pressure; FDP, fibrin degradation product; BUN, blood urea nitrogen.

Figure 8. Comparison of area under the curve using machine learning analysis with early features in the training set (A) and validation set (B)

(A) Training set



(B) Validation set

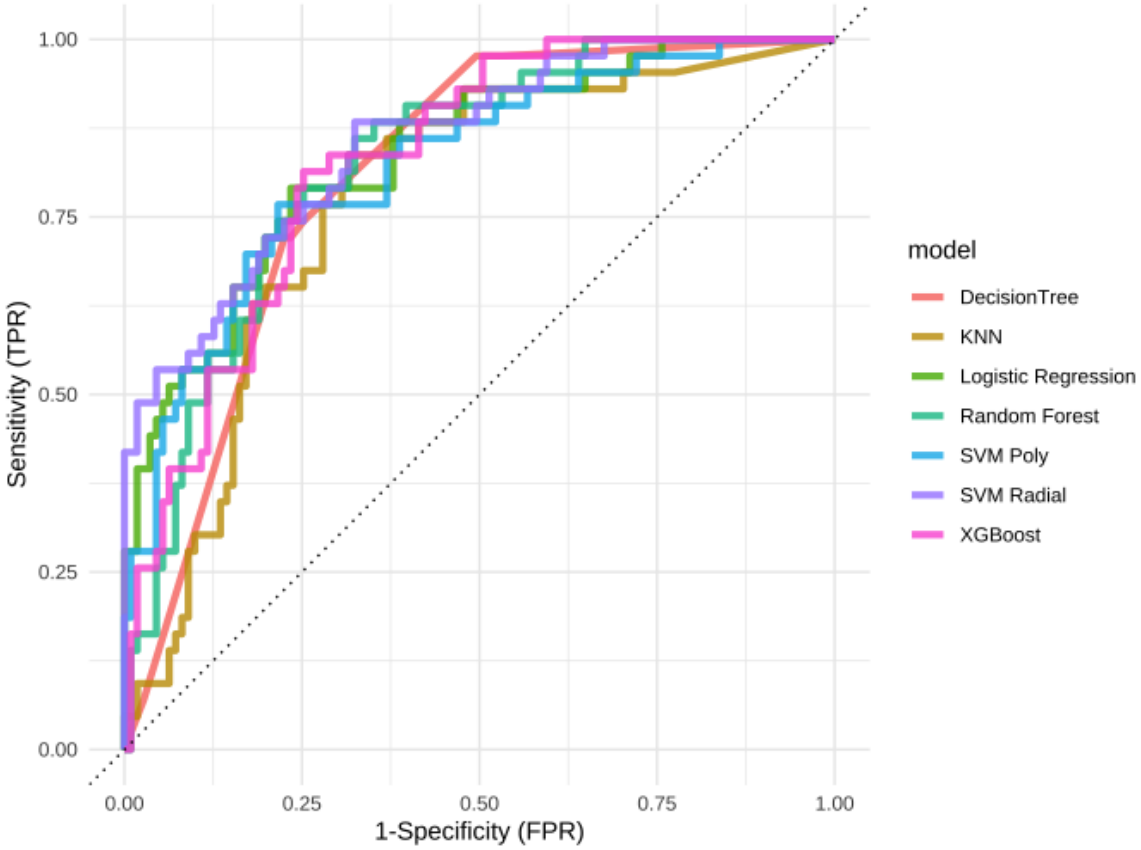
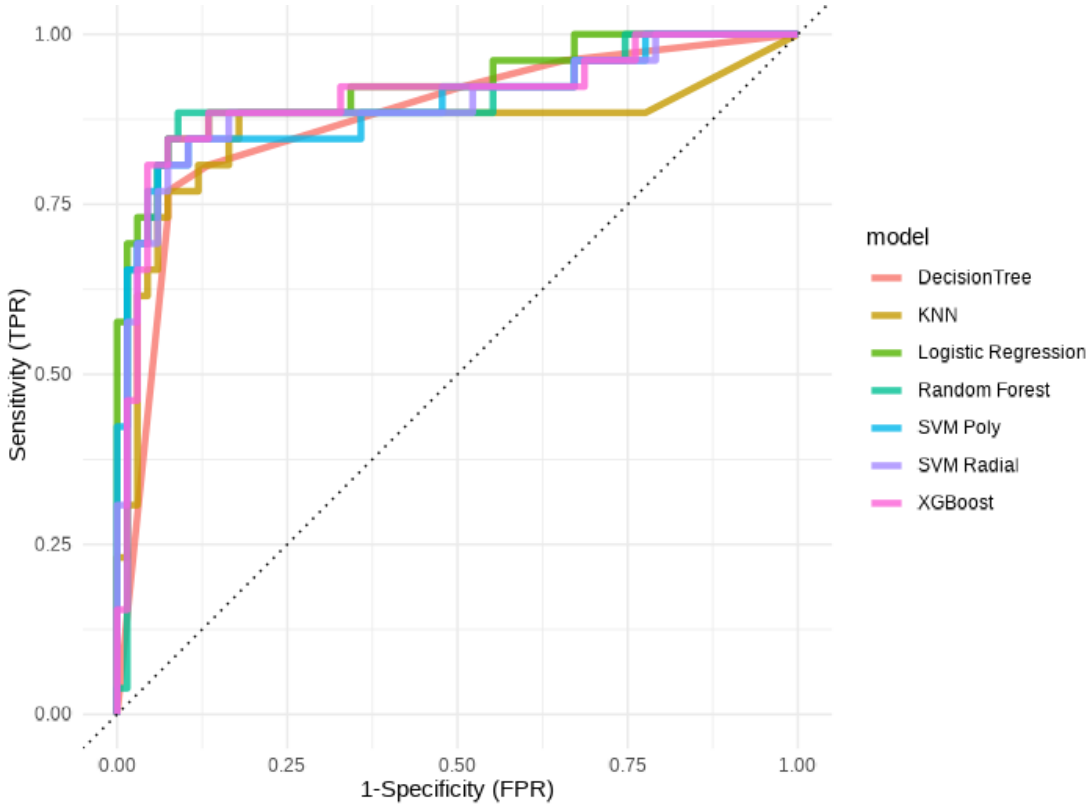


Figure 9. Comparison of area under the curve using machine learning analysis with all features in the training set (A) and validation set (B)

(A) Training set



(B) Validation set