RESEARCH ARTICLE

Data-driven predictive modeling for massive intraoperative blood loss during living donor liver transplantation: Integrating machine learning techniques

Taiichi Wakiya[1]*, Yukihiro Sanada[1], Noriki Okada[1], Yuta Hirata[1], Toshio Horiuchi[1], Takahiko Omameuda[1], Yasuharu Onishi[1], Yasunaru Sakuma[1], Hironori Yamaguchi[1], Yoshihiro Sasaki[2], Naohiro Sata[1]

1 Division of Gastroenterological, General and Transplant Surgery, Department of Surgery, Jichi Medical University, Shimotsuke, Tochigi, Japan, 2 Department of Medical Informatics, Hirosaki University Hospital, Hirosaki, Aomori, Japan

* wakiya.taiichi@jichi.ac.jp

Abstract

Background

Massive intraoperative bleeding (IBL) in liver transplantation (LT) poses serious risks and strains healthcare resources necessitating better predictive models for risk stratification. As traditional models often fail to capture the complex, non-linear patterns underlying bleeding risk, this study aimed to develop data-driven machine learning models for predicting massive IBL during living donor LT (LDLT) using preoperative factors.

Methods

Two hundred ninety consecutive LDLT cases from a prospective database were analyzed. Logistic regression models were built using 73 preoperative demographic and laboratory variables to predict massive IBL (≥ 80 mL/kg). The dataset was randomly split (70% training, 30% testing). The model was trained and validated through three-fold cross-validation, with backward stepwise feature selection iterated 100 times across unique random splits. The final model, based on a high stability index, was evaluated using the area under the curve (AUC).

Results

Massive IBL was observed in 141 patients (48.6%). In standard logistic regression, significant differences were found in 42 of 73 factors between groups stratified by massive IBL, however, substantial multicollinearity limited interpretability. In the feature selection across 100 iterations, the data-driven model achieved an average AUC of 0.840 in the validation and 0.738 in the test datasets. The final model, based on 11

selected features with a high stability index, achieved an AUC of 0.844. An easy-to-use online risk calculator for massive IBL was developed and is available at: https://tai1wakiya.shinyapps.io/ldlt_bleeding_ml/.

Conclusions

Our findings highlight the potential of machine learning in capturing complex risk factor interactions for predicting massive IBL in LDLT.

Introduction

Liver transplantation (LT) is a high-risk procedure frequently accompanied by substantial intraoperative bleeding (IBL), which can lead to adverse outcomes [1–3]. Additionally, perioperative bleeding and transfusions consume significant human and financial resources [4,5]. Accordingly, effective strategies to anticipate and reduce bleeding are essential for optimizing outcomes and maintaining the sustainability of transplant programs.

Despite ongoing advances, predicting critical IBL during LT remains a major clinical challenge. Risk stratification tools using conventional statistical methods, such as logistic regression, have been developed to identify patients at increased risk [6–9]. However, these models often assume linearity and independence among variables, limiting their ability to capture the complex, multifactorial nature of LT-associated bleeding. As a result, their accuracy and clinical utility remain suboptimal.

Machine learning offers a promising alternative by allowing the integration of numerous interrelated factors without requiring pre-specified assumptions about variable relationships. ML models can capture nonlinear associations and detect complex patterns that traditional approaches or clinical intuition may overlook. This capability also extends to handling high-dimensional datasets and identifying subtle but clinically relevant trends, potentially offering novel insights into perioperative risk [10,11]. This capability may enhance the predictive performance of bleeding models and facilitate more personalized perioperative management.

Although ML has shown value in various surgical contexts [12–14], its application in predicting bleeding outcomes in LT is still limited. Therefore, we aimed to develop a data-driven, ML-based model using prospectively collected preoperative variables to predict massive IBL in LDLT. Our objective was to establish a clinically applicable framework that outperforms conventional models and supports individualized surgical planning.

Materials and methods

Patients

We conducted a retrospective monocentric observational study using a prospectively maintained database. This study was approved by the Ethics Committee of Jichi Medical University (Approval No. 20−008). This study was designed and conducted in accordance with the principles of the Declaration of Helsinki and Istanbul. The need for written informed consent for the present study was waived by the

Institutional Review Board of Jichi Medical University in view of its retrospective design, in accordance with national and local guidelines, considering the fact that all clinical/laboratory measurements and procedures were part of routine care. This study included 290 consecutive patients who underwent LDLT at our facility between 2008 and 2024.

As an institutional policy, LDLT was not undertaken in the presence of any extrahepatic infection or when vasopressor or inotropic support was required. All recipients therefore proceeded to incision without active extrahepatic infection and without vasopressor or inotropic support.

Surgical procedures and operative management

Donor hepatectomy was selected based on the recipient's standard liver volume, weight, and graft volume determined using preoperative computed tomographic volumetry. For the recipient's operation, inverted T-shaped or transverse incisions were made, and a total hepatectomy was performed. The graft hepatic vein was anastomosed to the stump of the recipient's hepatic veins, which formed a single orifice, in an end-to-end manner. Hepatic artery reconstruction was routinely performed using microsurgical techniques. Choledocho-choledochostomy was the first choice for biliary reconstruction, except in cases in which the bile duct could not be used, such as with biliary atresia (BA) or primary sclerosing cholangitis.

Definition of massive intraoperative bleeding

At the start of surgery, pre-existing ascites was aspirated and excluded from IBL calculation. During the procedure, newly accumulated peritoneal fluid was included in the field in–out balance to calculate IBL. The volume was measured using suction canisters and swab weight, and it may include non-blood components to a variable extent, such as ascites, bile, or lymph. IBL was divided by body weight and referred to as adjusted intraoperative bleeding (aIBL, mL/kg). Massive aIBL was defined as aIBL ≥ 80 mL/kg. In clinical practice, the estimated circulating blood volume is approximately 70–80 mL/kg [15]. Therefore, aIBL exceeding 80 mL/kg is effectively equivalent to the loss of the entire circulating blood volume, indicating a clinically significant hemorrhagic event. Accordingly, 80 mL/kg serves as a physiologically meaningful and practical cutoff.

Data collection

For each patient, we collected demographic and laboratory data immediately before LT. The data were accessed for research purposes on April 22, 2025. Investigators had access to identifiable participant information during data collection; however, all data were de-identified prior to analysis. Seventy-three perioperative variables were extracted from the prospectively collected database. Patient demographic data included age, sex, body height, body weight, and etiology. Etiology was categorized as follows: acute liver failure (ALF), BA, graft failure, or other. Additionally, information regarding the history of previous abdominal surgery, ABO incompatibility, rituximab desensitization, and graft type was collected. Graft type was categorized as left lateral segment (LLS), left liver (LL), right liver (RL), right posterior segment (RPS), reduced LLS, or monosegment graft. To prevent statistical bias and enhance the generalizability of the model, incidence factors with a prevalence of less than 5%, such as spontaneous bacterial/fungal peritonitis and portal vein thrombosis, were excluded from the predictive variables. The 55 laboratory variables and their abbreviations are detailed in S1 Table. Prior to binary logistic regression analysis, the variables were standardized to have a mean of 0 and a standard deviation of 1, thereby ensuring that the logistic coefficients reflected the magnitude of influence on prediction. Given our focus on identifying modifiable risk factors, we placed particular emphasis on preoperative laboratory parameters, which can be influenced by conditions such as infection and thrombosis and can be actively managed before surgery. Accordingly, our candidate predictors were drawn primarily from preoperative indicators. In particular, we included pre-incision laboratory values because they integrate the cumulative effects of prior treatments, including the management of preoperative infections and the administration of blood products, and best capture the immediate preoperative physiological state. The study dataset used in this analysis is provided in S2 File.

Statistical analyses

Continuous variables were expressed as medians (ranges) and analyzed using nonparametric methods for non-normally distributed data (Mann–Whitney U test). Categorical variables were reported as numbers (percentages) and analyzed using the chi-square test or Fisher's exact test, as appropriate. Variables with a significant relationship to massive IBL in univariate analysis were used in a binary logistic regression model. The correlation between the two parameters was analyzed using the Spearman rank-order method. Differences were considered significant at $P < 0.05$. Statistical analyses were performed using GraphPad Prism (v10.2.3; GraphPad Software, San Diego, CA; USA, https://www.graphpad.com).

Data preparation and splitting

The overall workflow is illustrated using a block diagram in Fig 1. The dataset comprised 73 features along with corresponding binary labels. Data preprocessing involved a stratified random split, allocating 70% of the data for training and 30% for testing. To ensure a robust evaluation across different random splits, this process was repeated 100 times using unique random seeds (random_state = rr, where rr = 0, 1,..., 99).

Backward feature selection and model optimization

To identify the most informative predictors of massive intraoperative bleeding, we applied backward stepwise feature selection within a logistic regression framework. Model optimization and evaluation were conducted using stratified 3-fold cross-validation, repeated across 100 random data splits to ensure generalizability. At each iteration, variables were progressively removed based on their contribution to model performance, assessed using the area under the receiver operating characteristic curve (AUC). The feature set yielding the highest validation AUC was selected as the final model. Details of the feature elimination procedure, coefficient ranking, and model tuning are provided in S1 File. All analyses were conducted using Python libraries including Scikit-learn and Pandas [16]. Run instructions are provided in S3 File. The analysis script is provided in S4 File.

Development of an online calculator for estimating the risk of massive aIBL

We developed an online calculator based on the final logistic regression model to estimate the probability of massive aIBL. To ensure model robustness, feature stability was evaluated across multiple random data splits using occurrence rate and absolute mean weight. The summarization script is provided in S5 File. Features consistently selected with strong predictive weights were incorporated into the final model. Prior to application, selected features were standardized and weighted using their regression coefficients to calculate the predicted probability of massive aIBL. The calculator is available at: https://tai1wakiya.shinyapps.io/ldlt_bleeding_ml/. Details of the feature selection metrics, weighting procedures, and standardization process are provided in S1 File.

Results

Clinical significance of massive aIBL

Among the 290 patients, those with aIBL greater than 80 mL/kg exhibited significantly higher rates of graft loss compared to those with lower aIBL (Fig 2). These findings support the clinical validity of using 80 mL/kg as a reasonable and actionable threshold for defining massive aIBL in LDLT.

Comparison of patient characteristics in the massive aIBL and non-massive aIBL groups

The demographic data of the 290 enrolled patients are presented in Table 1, which also includes information on the IBL. Among all cases, massive aIBL was observed in 141 patients (48.6%). Except for indication and graft type, no significant differences were observed in the demographic data between the groups.
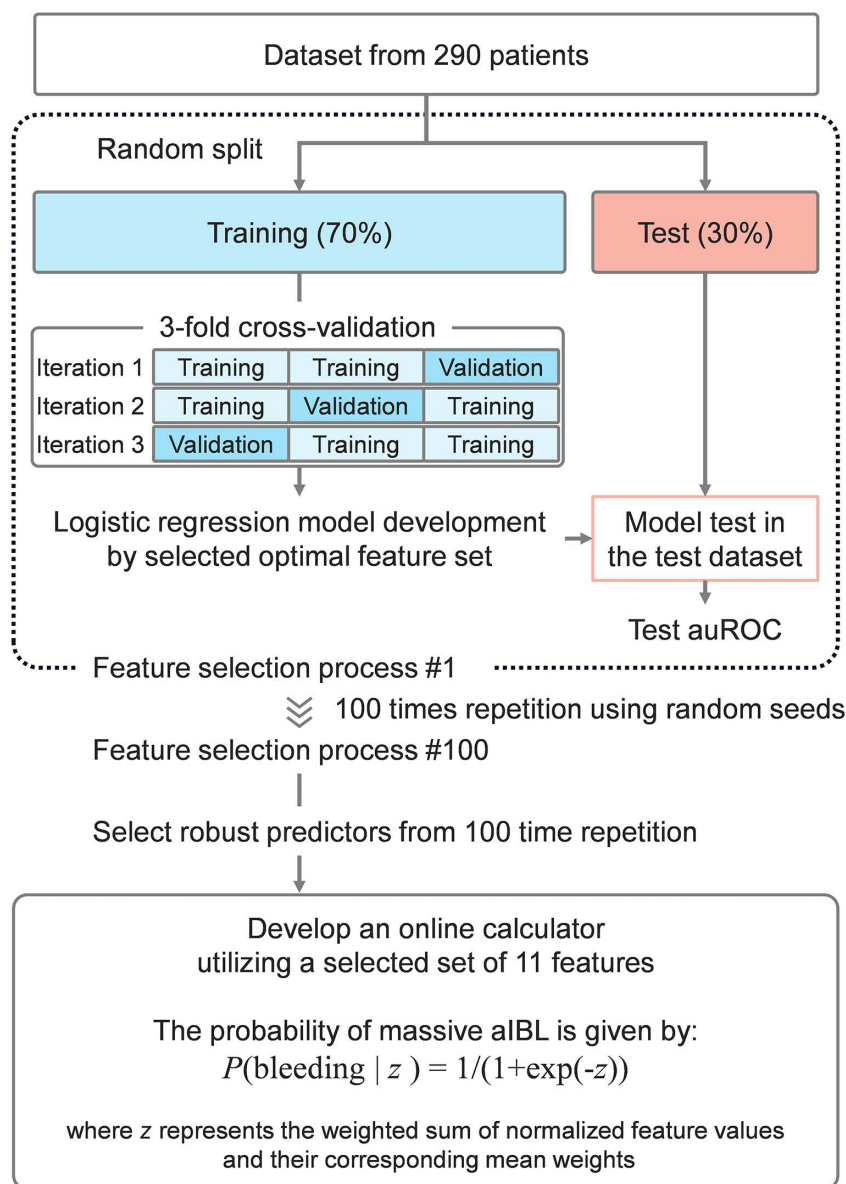
Fig 1. The study workflow and methodological process.

https://doi.org/10.1371/journal.pone.0326000.g001

## Comparison of laboratory data in the massive aIBL and non-massive aIBL groups

Laboratory data are presented in Table 2. Significant differences (P < 0.05) were observed in 40 of the 55 laboratory test items between the two groups. To predict the occurrence of massive aIBL, we conducted a binary logistic regression analysis with the occurrence of massive aIBL as the dependent variable. Forty significant predictor variables associated with massive aIBL, identified using univariate analysis (P < 0.05), were included in the binary logistic regression analysis. The analysis identified total bile acids (P = 0.013, odds ratio [OR] = 1.010, 95% confidence interval [CI]: 1.000–1.010), albumin (P = 0.014, OR = 0.228, 95% CI: 0.071–0.736), and ammonia (P = 0.043, OR = 1.020, 95% CI: 1.000–1.050) as significant predictors of incidences of massive aIBL. However, as noted in S2 Table, there was significant multicollinearity among
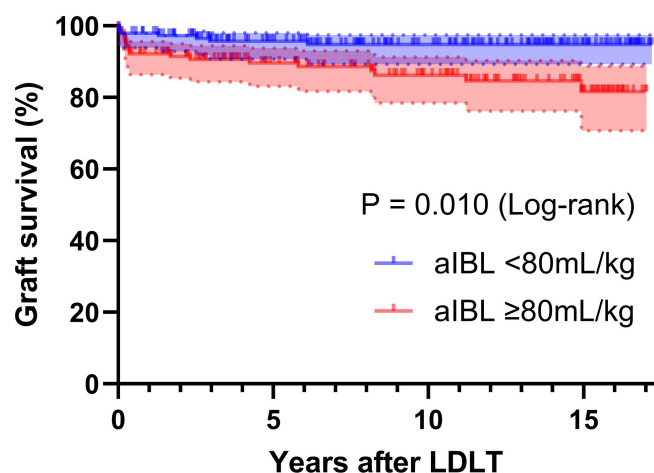
Fig 2. Graft survival curve stratified by adjusted intraoperative bleeding volume using an 80 mL/kg cutoff. IBL, intraoperative bleeding; LDLT, living donor liver transplantation.

https://doi.org/10.1371/journal.pone.0326000.g002

Table 1. Comparison of patient characteristics in the massive aIBL and non-massive aIBL groups.

	All (n = 290)	aIBL <80mL/kg (n = 149)	aIBL ≥80mL/kg (n = 141)	P value
Gender, female, n	168 (57.9)	88 (59.1)	80 (56.7)	0.778
Age, year	3 (0-69)	3 (0-65)	4 (0-69)	0.606
Body height, cm	94.9 (46.0-178.0)	95.6 (50.5-176.5)	93.0 (46.0-178.0)	0.480
Body weight, kg	14.2 (2.6-87.9)	14.1 (2.6-72.5)	14.4 (2.6-87.9)	0.673
Indication, n				0.046
ALF	14 (4.8)	3 (2.0)	11 (7.8)	
BA	155 (53.5)	83 (55.7)	72 (51.1)	
Graft failure	13 (4.5)	4 (2.7)	9 (6.4)	
Other	108 (37.2)	59 (39.6)	49 (34.8)	
History of previous abdominal surgery, n	198 (68.3)	98 (65.8)	100 (70.9)	0.415
ABO incompatible, n	58 (20.0)	27 (18.1)	31 (22.0)	0.499
Rituximab desensitization, n	59 (20.3)	34 (22.8)	25 (17.7)	0.352
Graft type, n				<0.001
LLS	127 (43.8)	78 (52.3)	49 (34.8)	
LL	74 (25.5)	43 (28.9)	31 (22.0)	
RL	57 (19.7)	20 (13.4)	37 (26.2)	
Monosegment	16 (5.5)	2 (1.3)	14 (9.9)	
Reduced LLS	13 (4.5)	5 (3.4)	8 (5.7)	
RPS	3 (1.0)	1 (0.7)	2 (1.4)	
IBL, mL	1012 (30-64510)	540 (30-5000)	2603 (250-64510)	<0.001
aIBL, mL/kg	78.6 (1.6-1112.4)	36.5 (1.6-79.5)	159.5 (80.4-1112.4)	<0.001

aIBL, adjusted intraoperative bleeding; ALF, acute liver failure; BA, biliary atresia; IBL, intraoperative bleeding; LL, left liver; LLS, left lateral segment; RL, right liver; RPS, right posterior segment.

https://doi.org/10.1371/journal.pone.0326000.t001

Table 2. Comparison of laboratory data in the massive aIBL and non-massive aIBL groups.

	All (n = 290)	aIBL <80mL/kg (n = 149)	aIBL ≥80mL/kg (n = 141)	P value
A2PI, %	73.5 (16.4-150.0)	86.1 (27.3-150.0)	60.7 (16.4-116.1)	<0.001
Alb, g/dL	3.0 (1.7-4.6)	3.2 (1.8-4.4)	2.8 (1.7-4.6)	<0.001
ALP, U/L	803.0 (119.0-7793.0)	857.0 (136.0-7452.0)	688.0 (119.0-7793.0)	0.320
ALT, U/L	49.5 (1.0-723.0)	50.0 (1.0-467.0)	48.0 (10.0-723.0)	0.633
AMY, U/L	43.0 (3.0-387.0)	45.0 (7.0-387.0)	38.0 (3.0-320.0)	0.027
AnGap, mmol/L	11.7 (2.0-26.4)	12.0 (3.4-26.4)	11.1 (2.0-19.6)	<0.001
AST, U/L	75.5 (14.0-1013.0)	62.0 (16.0-565.0)	88.0 (14.0-1013.0)	0.148
AT3, %	67.1 (17.1-157.5)	79.8 (21.8-157.5)	47.3 (17.1-138.5)	<0.001
BE, mmol/L	−1.2 (−9.5-12.0)	−1.4 (−9.5-8.9)	−0.7 (−8.0-12.0)	0.063
BS, mg/dL	108.0 (51.0-349.0)	111.0 (61.0-244.0)	104.0 (51.0-349.0)	0.024
BUN, mg/dL	8.0 (1.0-66.0)	7.0 (1.0-66.0)	9.0 (1.0-58.0)	<0.001
Ca, mg/dL	8.7 (7.4-12.4)	8.8 (7.4-12.4)	8.7 (7.5-12.4)	0.006
Ca2calc, mmol/L	1.1 (0.7-1.4)	1.1 (0.8-1.4)	1.1 (0.7-1.4)	0.417
Che, U/L	162.0 (25.0-443.0)	176.0 (25.0-443.0)	130.0 (33.0-436.0)	<0.001
Cl, mmol/L	107.0 (89.0-115.00	108.0 (90.0-115.0)	106.0 (89.0-114.0)	<0.001
CPK, U/L	65.0 (10.0-981.0)	67.0 (11.0-566.0)	64.0 (10.0-981.0)	0.624
Crea, mg/dL	0.24 (0.02-3.34)	0.23 (0.02-1.79)	0.24 (0.03-3.34)	0.191
CRP, mg/dL	0.20 (0.01-7.59)	0.13 (0.01-7.59)	0.51 (0.01-6.81)	<0.001
Dbil, mg/dL	0.68 (0.01-39.06)	0.28 (0.01-19.13)	2.24 (0.01-39.06)	<0.001
Ddimer, μg/mL	1.1 (0.1-47.8)	0.5 (0.1-16.6)	2.4 (0.1-47.8)	<0.001
Ferri, ng/mL	42.6 (3.0-3526.1)	28.3 (3.0-2370.0)	61.3 (4.1-3526.1)	<0.001
Fib, mg/dL	173.5 (42.0-592.0)	190.0 (42.0-592.0)	166.0 (43.0-356.0)	0.003
GGT, U/L	82.5 (8.0-1517.0)	112.0 (8.0-1517.0)	58.0 (8.0-1255.0)	0.529
Hb, g/dL	9.1 (4.1-13.5)	9.6 (6.4-13.5)	8.6 (4.1-13.4)	<0.001
HCO3	23.2 (16.8-36.5)	22.8 (16.8-33.0)	23.9 (17.6-36.5)	0.004
Hct, %	27.1 (11.3-39.9)	29.0 (19.8-39.9)	25.7 (11.3-38.5)	<0.001
K, mmol/L	4.2 (2.1-8.0)	4.1 (2.7-8.0)	4.3 (2.1-6.2)	0.340
Lactate, mmol/L	1.2 (0.4-10.4)	1.2 (0.4-10.4)	1.2 (0.5-8.5)	0.549
LDH, U/L	219.0 (98.0-938.0)	217.0 (98.0-938.0)	230.0 (101.0-754.0)	0.081
MCH, pg	29.7 (19.9-41.1)	29.1 (19.9-39.1)	30.7 (23.0-41.1)	<0.001
MCHC, g/dL	33.4 (30.0-36.1)	33.4 (30.0-36.1)	33.6 (30.3-36.0)	0.036
MCV, fL	88.5 (65.0-118.6)	86.0 (65.0-111.6)	91.0 (72.6-118.6)	<0.001
Mg, mg/dL	2.0 (0.8-2.9)	2.1 (1.3-2.5)	2.0 (0.8-2.9)	0.036
Na, mmol/L	139.0 (123.0-148.0)	140.0 (123.0-148.0)	138.0 (123.0-147.0)	<0.001
NH3, μmol/L	43.5 (4.0-161.0)	39.0 (4.0-119.0)	49.0 (15.0-161.0)	<0.001
P, mg/dL	4.1 (1.9-6.7)	4.5 (2.0-6.4)	3.8 (1.9-6.7)	<0.001
PCO2, mmHg	38.6 (28.3-102.0)	37.6 (29.0-54.8)	39.4 (28.3-102.0)	0.003
PH	7.39 (7.13-7.54)	7.40 (7.19-7.54)	7.39 (7.13-7.53)	0.887
PIC, μg/mL	0.5 (0.1-8.9)	0.4 (0.1-2.9)	0.6 (0.1-8.9)	<0.001
Plasminogen, %	62.1 (17.9-148.5)	71.0 (23.2-148.5)	54.2 (17.9-141.0)	<0.001
Plt, x 10⁴/μL	10.5 (1.4-66.5)	12.6 (2.4-66.5)	8.6 (1.4-34.9)	<0.001
PO2, mmHg	181.6 (69.4-463.0)	191.0 (93.3-463.0)	166.2 (69.4-386.3)	<0.001
ProteinC, %	44.5 (10.8-180.8)	57.9 (12.0-180.8)	34.0 (10.8-131.7)	<0.001
PTINR	1.35 (0.91-5.80)	1.25 (0.91-2.72)	1.47 (0.99-5.80)	<0.001

(Continued)

Table 2. (Continued)

	All (n = 290)	aIBL <80mL/kg (n = 149)	aIBL ≥80mL/kg (n = 141)	P value
PTpercent, %	57.7 (8.8-115.1)	68.4 (19.1-115.1)	49.7 (8.8-95.6)	<0.001
RBC, x 10⁴/μL	310.5 (122.0-461.0)	334.0 (210.0-461.0)	278.0 (122.0-442.0)	<0.001
TAT, ng/mL	2.5 (0.5-218.1)	2.0 (0.5-53.1)	3.0 (0.6-218.1)	<0.001
TBA, μmol/L	67.2 (0.1-619.6)	30.9 (0.1-493.6)	119.5 (0.4-619.6)	<0.001
Tbil, mg/dL	2.74 (0.11-50.01)	1.56 (0.11-25.63)	6.18 (0.16-50.01)	<0.001
Tchol, mg/dL	142.0 (46.0-1999.0)	149.0 (46.0-1999.0)	134.0 (56.0-1115.0)	0.010
TCO2, mmol/L	24.5 (17.8-38.3)	23.9 (17.8-34.4)	25.2 (18.6-38.3)	0.002
TG, mg/dL	49.0 (10.0-370.0)	49.0 (15.0-370.0)	49.0 (10.0-342.0)	0.499
TP, g/dL	5.7 (3.2-8.4)	5.8 (4.3-7.3)	5.6 (3.2-8.4)	0.113
UA, mg/dL	4.0 (0.6-9.9)	4.1 (0.7-9.5)	3.9 (0.6-9.9)	0.027
WBC, x 10³/μL	3.1 (0.5-15.7)	3.4 (0.6-10.2)	3.0 (0.5-15.7)	0.251

aIBL, adjusted intraoperative bleeding; IBL, intraoperative bleeding.

Abbreviations for laboratory test items are detailed in S1 Table.

many factors in this analysis. The presence of multicollinearity can lead to instability in estimates, wider confidence intervals, and increased errors in coefficient estimates, complicating interpretation and potentially leading to incorrect conclusions. These results indicate that to construct a predictive model that exhibits multicollinearity with our dataset, it is crucial to select variables objectively, excluding subjectivity, and to utilize alternative analytical methods.
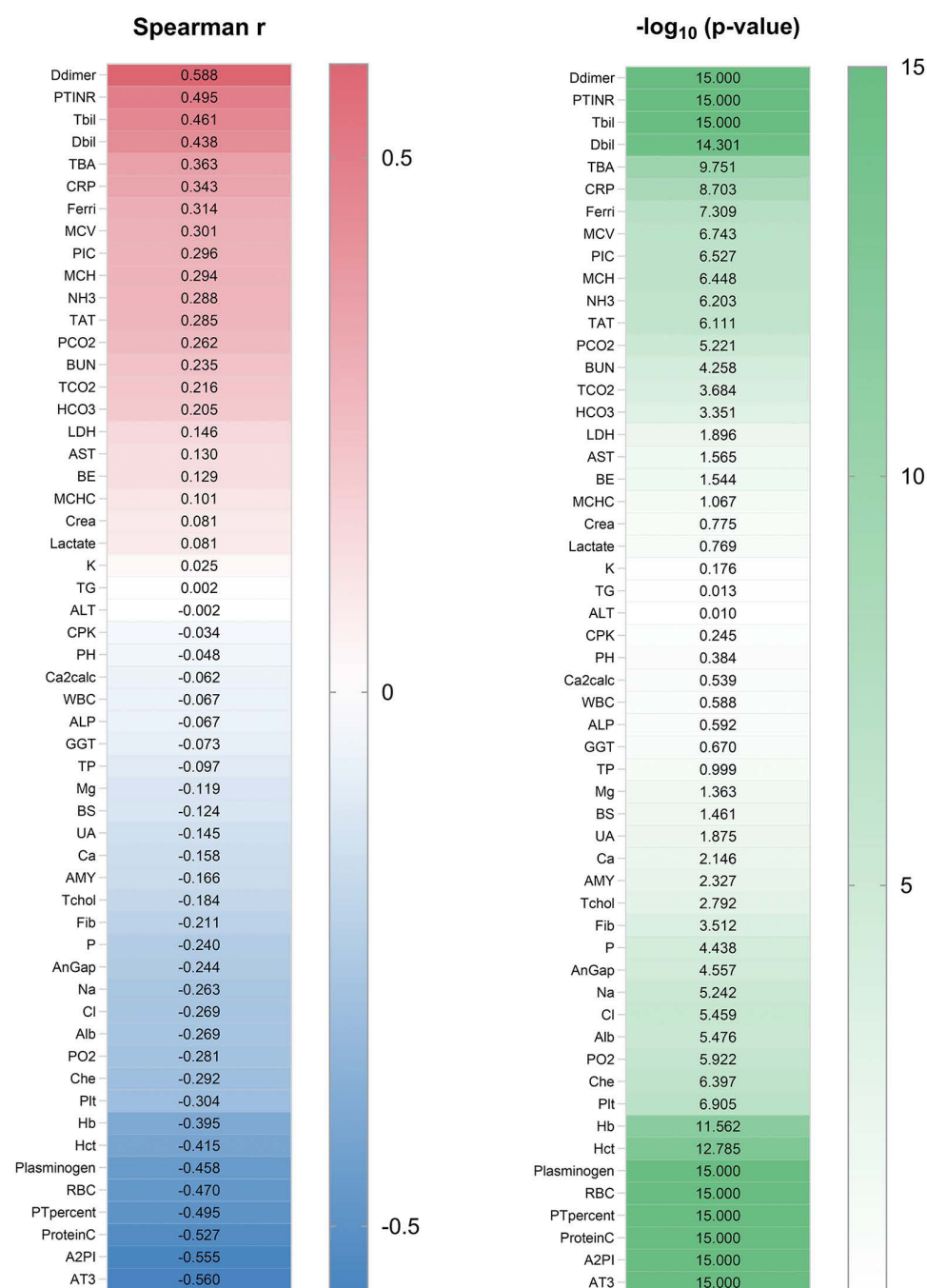
Correlations between preoperative laboratory tests and intraoperative bleeding volume

Next, we examined the relationship between the preoperative laboratory test results and aIBL. Of the 55 parameters tested, 42 showed significant correlations with the bleeding volume. Positive correlations were found for 43.6% of the parameters, whereas 56.4% exhibited negative correlations. The top five positively correlated biomarkers were D-dimer (r = 0.588), prothrombin time international normalized ratio (r = 0.495), total bilirubin (r = 0.461), direct bilirubin (r = 0.438), and total bile acids (r = 0.363). Conversely, antithrombin III (r = −0.560), alpha 2-plasmin inhibitor (r = −0.555), protein C (r = −0.527), prothrombin time percent (r = −0.495), and plasminogen (r = −0.458) were highly negatively correlated (Fig 3). These findings suggest that these markers can potentially guide preoperative assessment and intervention.

Feature selection and model performance in logistic regression for massive aIBL prediction

We further explored a more robust prediction model for massive aIBL. We used the backward stepwise method for feature selection to construct our binary logistic regression model. During this process, we calculated the absolute magnitudes of the coefficients and removed the least influential covariates. This analysis incorporated laboratory variables as well as demographic data, including age, categorical age (under 6 years, 6–18 years, or 18 years and older), sex, body height, body weight, etiologies (ALF, BA, graft failure, or other), history of previous abdominal surgery, ABO incompatibility, rituximab desensitization, and graft type (LLS, LL, RL, RPS, reduced LLS, or monosegment). We started with all 73 features and systematically removed the least significant ones from the set. One of the representative feature selection processes is presented in S3 Table and S1 Fig. The process for narrowing down the variables is detailed in S3 Table. Antithrombin III was identified as the final covariate in the dataset. S1 Fig illustrates the performance of the model at each stage. The highest validation AUC of 0.867 was observed at the stage with 12 covariates. The AUC obtained in the test dataset was 0.728. To ensure a robust evaluation, this process was repeated 100 times across different random splits (S2 Fig). The

Fig 3. Correlations between preoperative laboratory tests and intraoperative bleeding volume. The left panel shows Spearman's rank correlation coefficients, and the right panel shows p-values. All abbreviations and their units are listed in S1 Table.

https://doi.org/10.1371/journal.pone.0326000.g003

average number of selected features across these 100 trials was 15.1. The average AUC in the validation and test datasets was 0.840 (standard deviation [SD] = 0.022) and 0.738 (SD = 0.046), respectively. These findings support a systematic approach for preoperatively predicting massive IBL in LDLT, suggesting the robustness and reliability of feature selection and model performance.

Identifying key predictors and constructing an online calculator for massive aIBL

To evaluate the consistency of selected features across 100 random data splits, index stability was assessed using occurrence rate and absolute mean weight (Fig 4). Based on the results from 100 iterations, features meeting the predefined criteria of an occurrence rate of at least 50% and an absolute mean weight of 0.5 or greater were selected. A total of 11 features were identified as robust predictors. Among them, six features exhibited positive mean weights, indicating a direct association with massive aIBL occurrence: D-dimer, ferritin, age, total bile acids, gamma-glutamyl transferase, and ammonia. Conversely, five features demonstrated negative mean weights, suggesting a potential protective effect: fibrinogen, platelet count, albumin, LLS graft, and antithrombin III. To enhance clinical applicability and ensure streamlined implementation, we developed a final prediction model for massive aIBL using these 11 selected features. In the test across all 290 cases, the model achieved an AUC of 0.844, with a sensitivity of 74.5%, specificity of 78.5%, positive predictive value of 76.6%, and negative predictive value of 76.5%. To further facilitate clinical adoption, we developed a user-friendly online calculator based on this risk model and made available online at: https://tai1wakiya.shinyapps.io/ldlt_bleeding_ml/.

Discussion

In this study, we developed and validated ML models to predict massive IBL during LDLT. Using preoperative demographic and laboratory data, we employed logistic regression models with backward feature selection to achieve high predictive accuracy. Furthermore, we developed an online calculator that provides physicians with an accessible tool to estimate an individualized risk of massive IBL in the preoperative setting. To our knowledge, this is the first study to apply a ML model for IBL prediction in LT and to translate these findings into a clinically applicable prediction tool. The true innovation of this study lies not only in the excellent accuracy of the model but also in showcasing a groundbreaking artificial intelligence-based approach with significant future potential for management in LT.
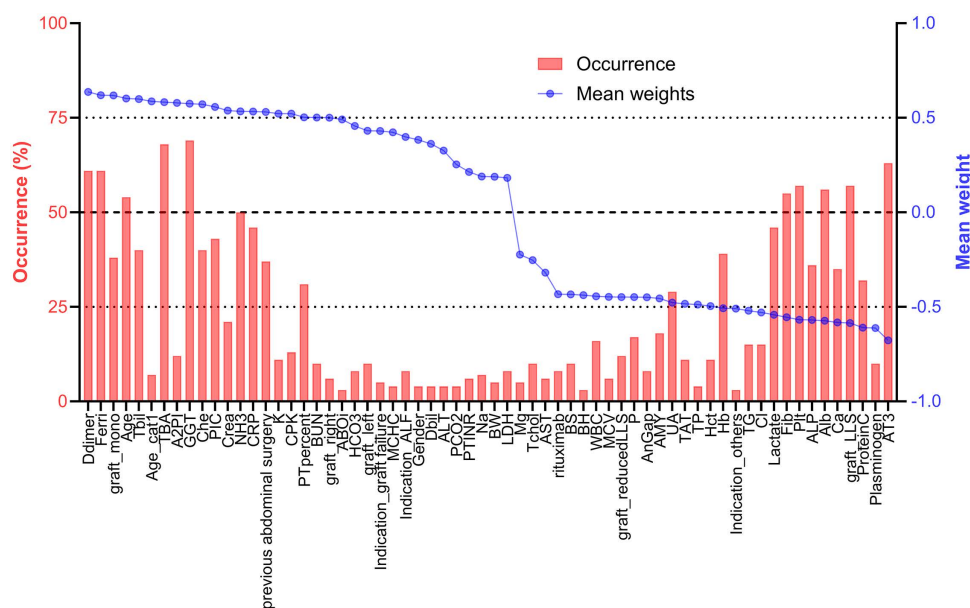
Fig 4. Index stability across 100 iterations. The index is arranged in order of occurrence, sorted by mean weight. All abbreviations and their units are listed in S1 Table.

https://doi.org/10.1371/journal.pone.0326000.g004

Previous studies have attempted to predict bleeding and transfusion requirements in LT by using various clinical parameters and traditional standard statistical models [6–9]. While these studies have provided valuable insights, their models often exhibit suboptimal predictive performance and limited clinical applicability. Standard methods in clinical studies have several limitations, including the selection of variables, confounding factors, and multicollinearity, as demonstrated in this study. Even in prospective studies, the issue of multicollinearity remains a challenge to resolve. Furthermore, standard methods often fail to capture the complex, non-linear relationships between variables that significantly influence bleeding risk. In contrast, our study goes beyond traditional analyses by employing a data-driven approach, allowing us to systematically identify not only well-established risk factors but also previously unrecognized predictors.

Our study leveraged ML techniques, which are used to identify intricate patterns and interactions among multiple factors, demonstrating superior accuracy and robustness in predicting massive IBL. ML offers several key benefits, including increased accuracy through the simultaneous analysis of multiple variables and their interrelationships, and the ability to identify complex patterns and correlations that traditional methods might overlook. By incorporating a comprehensive set of preoperative variables and employing advanced feature selection methods, our ML models offer a more reliable and practical tool for preoperative risk stratification in LDLT. These models effectively manage complex, high-dimensional datasets, uncover hidden correlations, and explore non-linear medical relationships, enhancing predictive capabilities and potentially leading to new scientific insights.

Crucially, we went beyond model development by translating these findings into a tangible clinical tool. To ensure practical utility in real-world clinical settings, we developed an online calculator based on our predictive model, providing an accessible platform for physicians to rapidly assess massive IBL risk. This helps bridge the gap between advanced computational modeling and clinical decision-making, while integrating ML into clinical practice, which may contribute to improving patient outcomes and optimizing resource management.

Antithrombin III and ferritin emerged as significant predictors in our machine learning-based model, providing new insights into bleeding risk in LT. Antithrombin, a liver-derived serine protease inhibitor, is a key physiological anticoagulant with additional anti-inflammatory effects mediated through endothelial interactions [17]. In cardiovascular surgery, low preoperative antithrombin III levels have been consistently associated with increased intraoperative blood loss [18]. Mechanistically, reduced antithrombin activity allows excess thrombin and factor Xa activity, promotes systemic microthrombosis and a hypercoagulable state, and accelerates the consumption of coagulation factors, platelets, and antithrombin itself. Secondary fibrinolysis then ensues, leading to a bleeding tendency. Antithrombin supplementation can, in principle, interrupt this cycle by suppressing thrombin generation and indirectly support hemostasis through endothelial protection. However, clinical outcomes have been inconsistent across studies and meta-analyses [19–22]. Similarly, although the available studies are both dated and limited in number, evidence supporting a positive effect of antithrombin III supplementation in LT remains limited [23,24]. Therefore, while antithrombin appears useful as a biomarker for risk stratification, effective intervention in the LT setting will likely require targeting the early hypercoagulable phase with optimized dosing and timing rather than antithrombin restoration alone.

Regarding ferritin, no direct evidence has previously established its role in predicting bleeding in liver surgery, including LT. Ferritin reflects both iron metabolism and systemic inflammation, and its elevation may indicate a preoperative inflammatory state that predisposes patients to coagulopathy [25]. Iron overload induces oxidative stress through Fenton chemistry, which causes endothelial dysfunction and consequently promotes bleeding [26]. Inflammatory responses are known to disrupt hemostasis through mechanisms such as endothelial dysfunction, hyperfibrinolysis, and altered coagulation cascades [27]. Supporting this, our analysis identified a strong positive correlation between IBL and C-reactive protein, another established inflammatory marker; however, ferritin demonstrated greater stability as a predictive index than C-reactive protein. Together, these findings illustrate how machine learning can reveal biologically and clinically relevant predictors that are often overlooked by conventional approaches. Ferritin appears to be a useful biomarker for risk stratification, but interventional evidence is limited. No randomized controlled trial has shown that lowering ferritin through iron

chelation reduces IBL in LT. Therefore, the benefit of iron-metabolism–targeted hemostatic strategies remains uncertain and warrants prospective evaluation.

This model was developed and calibrated in an LDLT cohort and is therefore positioned primarily for LDLT work-flows; application to orthotopic liver transplantation (OLT) should be undertaken with caution. In settings where LDLT predominates, including Japan, OLT often represents a different case mix, with a higher prevalence of ALF and higher MELD scores, which may shift the relative importance of predictors. As a small exploratory check, we applied the LDLT model without recalibration to our 13 OLT recipients treated during the study period and obtained 9 of 13 correct classifications, corresponding to sensitivity 75%, specificity 60%, accuracy 69%, and AUC 0.694 (S4 Table). While this suggests some portability, broader applicability will require model updating on dedicated OLT datasets and, ideally, training and validation in mixed multicenter cohorts that include both OLT and LDLT.

The findings of this current study should be interpreted in light of several limitations. First, because fluid that accumulates intraoperatively can contain non-blood components, our measurement may slightly overestimate true blood loss. Although pre-existing ascites at laparotomy was excluded from IBL measurement, complete separation of blood from newly accu-mulated non-blood peritoneal fluid during surgery is difficult in routine practice. In anticipation of multicenter validation, standardizing IBL definitions and measurement across institutions would improve comparability and model transportability. Second, some clinically relevant factors could not be incorporated. For example, viscoelastic testing with ROTEM, which informs coagulopathy management, was not routinely available during the study period. These variables were therefore not included in feature selection. The candidate feature set warrants reconsideration in future work, and incorporating such additional features may further improve model discrimination and calibration. Third, it was a single-institution cohort study with a relatively small patient population, which limits generalizability. Additional training data could potentially enhance the prediction accuracy. Furthermore, the lack of external validation using an independent dataset is also a limitation. One rea-son for the absence of external validation is the argument made by some biostatistics experts in predictive research stating that independent verification can be misleading and should be omitted as a model evaluation step [28,29]. These experts report that simulations confirming at least 100 events and 100 non-events are required for a reliable assessment of predic-tive performance. They suggest using all available data for model development, with some form of cross-validation or boot-strap validation to assess the statistical optimism in average predictive performance [29,30]. Based on these biostatistical perspectives, we chose to build our ML model using all data with cross-validation, performing 100 iterations with randomly split datasets to ensure robustness and generalizability. However, we acknowledge that the results of predictive research with small sample sizes are exploratory in nature [29]. Nevertheless, external validation in various clinical settings, cover-ing the heterogeneity among cases, is essential for clinical application. Given the high prediction accuracy of our method, further development using large databases, such as national or regional datasets, is expected and necessary.

Conclusions

Our study demonstrates the potential of ML to predict massive IBL during LDLT. These data-driven predictive models could revolutionize preoperative planning and intraoperative management, ultimately improving patient care and resource efficiency. As some predictors are exploratory, they should be interpreted cautiously until externally validated. Future research should focus on external validation using large datasets and the integration of these models into clinical practice.

Supporting information

S1 Fig. AUC by step in backward selection. AUC at each step of backward stepwise feature selection for the binary logistic regression model, shown for one representative iteration (one of 100 repetitions). AUC, area under the receiver operating characteristic curve.
(TIF)

S2 Fig. Best performance across 100 runs. Model performance from backward feature selection repeated 100 times; the best-performing result is shown for each random split. AUC, area under the receiver operating characteristic curve. (TIF)

S1 Table. List of abbreviations.
(DOCX)

S2 Table. List of variance inflation factors.
(DOCX)

S3 Table. Backward stepwise variable reduction.
(DOCX)

S4 Table. Performance in OLT cases.
(DOCX)

S1 File. Supplementary methods.
(DOCX)

S2 File. Dataset used in this study.
(TXT)

S3 File. Technical appendix – reproducibility and run guide.
(TXT)

S4 File. Source code 1 – model training and validation.
(PY)

S5 File. Source code 2 – feature selection and stability.
(PY)

Author contributions

Conceptualization: Taiichi Wakiya.

Data curation: Taiichi Wakiya, Yukihiro Sanada, Noriki Okada, Yuta Hirata, Toshio Horiuchi, Takahiko Omameuda, Yasuharu Onishi, Yasunaru Sakuma, Hironori Yamaguchi, Naohiro Sata.

Formal analysis: Taiichi Wakiya, Yoshihiro Sasaki.

Visualization: Taiichi Wakiya.

Writing – original draft: Taiichi Wakiya.

Writing – review & editing: Yukihiro Sanada, Noriki Okada, Yuta Hirata, Toshio Horiuchi, Takahiko Omameuda, Yasuharu Onishi, Yasunaru Sakuma, Hironori Yamaguchi, Yoshihiro Sasaki, Naohiro Sata.

References

1. Rana A, Petrowsky H, Hong JC, Agopian VG, Kaldas FM, Farmer D, et al. Blood transfusion requirement during liver transplantation is an important risk factor for mortality. J Am Coll Surg. 2013;216(5):902–7. https://doi.org/10.1016/j.jamcollsurg.2012.12.047 PMID: 23478547

2. Yuasa T, Niwa N, Kimura S, Tsuji H, Yurugi K, Egawa H, et al. Intraoperative blood loss during living donor liver transplantation: an analysis of 635 recipients at a single center. Transfusion. 2005;45(6):879–84. https://doi.org/10.1111/j.1537-2995.2005.04330.x PMID: 15934985

3. Ramos E, Dalmau A, Sabate A, Lama C, Llado L, Figueras J, et al. Intraoperative red blood cell transfusion in liver transplantation: influence on patient outcome, prediction of requirements, and measures to reduce them. Liver Transpl. 2003;9(12):1320–7. https://doi.org/10.1016/jlts.2003.50204 PMID: 14625833

4. Stokes EA, Wordsworth S, Staves J, Mundy N, Skelly J, Radford K, et al. Accurate costs of blood transfusion: a microcosting of administering blood products in the United Kingdom National Health Service. Transfusion. 2018;58(4):846–53. https://doi.org/10.1111/trf.14493 PMID: 29380872

5. Ruiz J, Dugan A, Davenport DL, Gedaly R. Blood transfusion is a critical determinant of resource utilization and total hospital cost in liver transplantation. Clin Transplant. 2018;32(2):10.1111/ctr.13164. https://doi.org/10.1111/ctr.13164 PMID: 29193383

6. Priem F, Karakiewicz PI, McCormack M, Thibeault L, Massicotte L. Validation of 5 models predicting transfusion, bleeding, and mortality in liver transplantation: an observational cohort study. HPB (Oxford). 2022;24(8):1305–15. https://doi.org/10.1016/j.hpb.2022.01.002 PMID: 35131142

7. Giabicani M, Joly P, Sigaut S, Timsit C, Devauchelle P, Dondero F, et al. Predictive role of hepatic venous pressure gradient in bleeding events among patients with cirrhosis undergoing orthotopic liver transplantation. JHEP Rep. 2024;6(6):101051. https://doi.org/10.1016/j.jhepr.2024.101051 PMID: 38699073

8. Thibeault F, Plourde G, Fellouah M, Ziegler D, Carrier FM. Preoperative fibrinogen level and blood transfusions in liver transplantation: A systematic review. Transplantation Reviews. 2023;37(4):100797. https://doi.org/10.1016/j.trre.2023.100797

9. Pustavoitau A, Rizkalla NA, Perlstein B, Ariyo P, Latif A, Villamayor AJ, et al. Validation of predictive models identifying patients at risk for massive transfusion during liver transplantation and their potential impact on blood bank resource utilization. Transfusion. 2020;60(11):2565–80. https://doi.org/10.1111/trf.16019 PMID: 32920876

10. Ray S, editor A Quick Review of Machine Learning Algorithms. 2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon); 2019 14-16 Feb. 2019.

11. Xu Y, Liu X, Cao X, Huang C, Liu E, Qian S, et al. Artificial intelligence: A powerful paradigm for scientific research. Innovation (Camb). 2021;2(4):100179. https://doi.org/10.1016/j.xinn.2021.100179 PMID: 34877560

12. Banerjee A, Dashtban A, Chen S, Pasea L, Thygesen JH, Fatemifar G, et al. Identifying subtypes of heart failure from three electronic health record sources with machine learning: an external, prognostic, and genetic validation study. Lancet Digit Health. 2023;5(6):e370–9. https://doi.org/10.1016/S2589-7500(23)00065-1 PMID: 37236697

13. Taleb I, Kyriakopoulos CP, Fong R, Ijaz N, Demertzis Z, Sideris K, et al. Machine Learning Multicenter Risk Model to Predict Right Ventricular Failure After Mechanical Circulatory Support: The STOP-RVF Score. JAMA Cardiol. 2024;9(3):272–82. https://doi.org/10.1001/jamacardio.2023.5372 PMID: 38294795

14. Wakiya T, Ishido K, Kimura N, Nagase H, Kubota S, Fujita H, et al. Prediction of massive bleeding in pancreatic surgery based on preoperative patient characteristics using a decision tree. PLoS One. 2021;16(11):e0259682. https://doi.org/10.1371/journal.pone.0259682 PMID: 34752505

15. Wilson F. Blood volume maintenance and restoration. In: Burton VW, Davies AH, Kilpatrick A, McIllmurray MB, Pring JE, Wilson F, editors. Essential Accident and Emergency Care. Dordrecht: Springer Netherlands; 1981. p. 130–40.

16. Pedregosa F. Scikit-learn: Machine learning in python. J Machine Learn Res. 2011;12:2825–30.

17. Rodgers GM, Mahajerin A. Antithrombin Therapy: Current State and Future Outlook. Clin Appl Thromb Hemost. 2023;29:10760296231205279. https://doi.org/10.1177/10760296231205279 PMID: 37822179

18. Laveson JE, Winford J, Iden J, Marengo-Rowe AJ, Race GJ. Investigation Of The Relationship Between Preoperative Chromogenic Antithrombin III Assays And Hemorrhage Following Heart Surgery. Thromb Haemost. 1981;46(05):0937. https://doi.org/10.1055/s-0038-1652889

19. Paparella D, Rotunno C, De Palo M, Finamore S, Guida P, Rubino G, et al. Antithrombin administration in patients with low antithrombin values after cardiac surgery: a randomized controlled trial. Ann Thorac Surg. 2014;97(4):1207–13. https://doi.org/10.1016/j.athoracsur.2013.11.040 PMID: 24507941

20. Jooste EH, Scholl R, Wu Y-H, Jaquiss RDB, Lodge AJ, Ames WA, et al. Double-Blind, Randomized, Placebo-Controlled Trial Comparing the Effects of Antithrombin Versus Placebo on the Coagulation System in Infants with Low Antithrombin Undergoing Congenital Cardiac Surgery. J Cardiothorac Vasc Anesth. 2019;33(2):396–402. https://doi.org/10.1053/j.jvca.2018.05.052 PMID: 30072263

21. Slaughter TF, Mark JB, El-Moalem H, Hayward KA, Hilton AK, Hodgins LP, et al. Hemostatic effects of antithrombin III supplementation during cardiac surgery: results of a prospective randomized investigation. Blood Coagul Fibrinolysis. 2001;12(1):25–31. https://doi.org/10.1097/00001721-200101000-00004 PMID: 11229823

22. Li T, Bo F, Meng X, Wang D, Ma J, Dai Z. The effect of perioperative antithrombin supplementation on blood conservation and postoperative complications after cardiopulmonary bypass surgery: A systematic review, meta-analysis and trial sequential analysis. Heliyon. 2023;9(11):e22266. https://doi.org/10.1016/j.heliyon.2023.e22266 PMID: 38053853

23. Baudo F, DeGasperi A, deCataldo F, Caimi TM, Cattaneo D, Redaelli R, et al. Antithrombin III supplementation during orthotopic liver transplantation in cirrhotic patients: a randomized trial. Thromb Res. 1992;68(4–5):409–16. https://doi.org/10.1016/0049-3848(92)90099-v PMID: 1290169

24. Palareti G, Legnani C, Maccaferri M, Gozzetti G, Mazziotti A, Martinelli G, et al. Coagulation and fibrinolysis in orthotopic liver transplantation: role of the recipient's disease and use of antithrombin III concentrates. S. Orsola Working Group on Liver Transplantation. Haemostasis. 1991;21(2):68–76. https://doi.org/10.1159/000216206 PMID: 1959800

25. Kell DB, Pretorius E. Serum ferritin is an important inflammatory disease marker, as it is mainly a leakage product from damaged cells. Metallomics. 2014;6(4):748–73. https://doi.org/10.1039/c3mt00347g PMID: 24549403

26. Dos Santos L, Bertoli SR, Ávila RA, Marques VB. Iron overload, oxidative stress and vascular dysfunction: Evidences from clinical studies and animal models. Biochim Biophys Acta Gen Subj. 2022;1866(9):130172. https://doi.org/10.1016/j.bbagen.2022.130172 PMID: 35597504

27. Iba T, Helms J, Neal MD, Levy JH. Mechanisms and management of the coagulopathy of trauma and sepsis: trauma-induced coagulopathy, sepsis-induced coagulopathy, and disseminated intravascular coagulation. J Thromb Haemost. 2023;21(12):3360–70. https://doi.org/10.1016/j.jtha.2023.05.028 PMID: 37722532

28. Steyerberg EW, Harrell FE Jr. Prediction models need appropriate internal, internal-external, and external validation. J Clin Epidemiol. 2016;69:245–7. https://doi.org/10.1016/j.jclinepi.2015.04.005 PMID: 25981519

29. Steyerberg EW. Validation in prediction research: the waste by data splitting. J Clin Epidemiol. 2018;103:131–3. https://doi.org/10.1016/j.jclinepi.2018.07.010 PMID: 30063954

30. Steyerberg EW, Harrell FE Jr, Borsboom GJ, Eijkemans MJ, Vergouwe Y, Habbema JD. Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. J Clin Epidemiol. 2001;54(8):774–81. https://doi.org/10.1016/s0895-4356(01)00341-9 PMID: 11470385