

Predicting the required pre-surgery blood volume in surgical patients based on machine learning

Ruilin Li^{1,3}, Xinyin Han^{1,3}, Liping Sun², Yannan Feng², Xiaolin Sun², Xiaoyu He^{1,3}, Yu Zhang^{1,3}, Haidong Zhu^{1,3}, Dan Zhao^{1,3}, Chuangchuang Dai^{1,3}, Zhipeng He^{1,3}, Shanyu Chen^{1,3}, Xin Wang⁶, Weizhong Li⁵, Xuebin Chi^{1,3,4}, Yang Yu^{2,*}, Beifang Niu^{1,3,*} and Deqing Wang^{2,*}

¹ Computer Network Information Center, Chinese Academy of Sciences, Beijing 100190, China

² Department of Blood Transfusion, Chinese PLA General Hospital, Beijing 100853, China

³ University of Chinese Academy of Sciences, Beijing 100190, China

⁴ Center of Scientific Computing Applications& Research, Chinese Academy of Sciences, Beijing 100190, China

⁵ J. Craig Venter Institute, La Jolla, CA 92037, USA

⁶ Pfizer (China) Research and Development Co., Ltd, Shanghai 201203, China

*Joint Corresponding Authors.

Deqing Wang, E-mail: deqingw@vip.sina.com, Department of Blood Transfusion, Chinese PLA General Hospital, No. 28, Fuxing Rd, Beijing 200853, China, Phone: 010-66936508, Fax: 010-66936508,

Beifang Niu, E-mail: niubf@cnic.cn, Computer Network Information Center, Chinese Academy of Sciences No.4 South Street, Zhongguancun, Haidian District, Beijing, China, Phone: 010-58812132, Fax: 010-58812114,

Yang Yu, E-mail: yuyangpla301@163.com, Department of Blood Transfusion, Chinese PLA General Hospital, No. 28, Fuxing Rd, Beijing 200853, China, Phone: 010-66935132, Fax: 010-66935132

Abstract Precisely predicting the required pre-surgery blood volume (PBV) in surgical patients is a formidable challenge in China. Inaccurate estimation is associated with excessive costs, postponed surgeries and adverse outcome after surgery due to insufficient supply or inventory. This study aimed to predict required PBV based on machine learning techniques. 181,027 medical documents over 6 years were cleaned and finally obtained 92,057 blood transfusion records. The blood transfusion and surgery related factors of perioperative patients, surgeons experience volumes and the actual volumes of transfused RBCs were extracted. 6 machine learning algorithms were used to build prediction models. The surgery patients received allogenic RBCs or without transfusion, had total volume less than 10 units, or had the latest laboratory examinations of pre-surgery within 7 days were included, providing 118,823 data points. 39 predictive factors related to the RBCs transfusion were identified. Random forest model was selected to predict the required PBV of RBCs with 72.9% accuracy and strikingly improved the accuracy by 30.4% compared with surgeons experience, where 90% of data was used for training. We tested and demonstrated that both the data-driven models and the random forest model achieved higher accuracy than surgeons experience. Furthermore, we developed a computational tool, PTRBC, to precisely estimate the required PBV in surgical patients and we believe this tool will find more applications in assisting clinician decisions, not only confined to making accurate pre-surgery blood requirement predicting.

Key words red blood cells, pre-surgery blood volume, random forest, prediction, modeling

1 Introduction

Insufficient supply of blood remains a challenge in many countries and this is particularly evident in China [1]. Electronic crossmatch has not been carried out yet and the vast majority of pre-surgery blood preparations are still based on serological crossmatch. The red blood cells (RBCs) transfusion is an important means to rectify anemia and improve oxygen supply. Almost 40 years ago, the decision to transfuse, the “transfusion trigger,” relied on factors that affect blood transfusion, such as hemoglobin (Hb), hematocrit (Hct), fatigue, and anemia symptoms [2]. The optimal RBCs transfusion is also still a focus of controversy, and clinicians mainly rely on experience [3]. For the patients with a selective surgery, the overestimate of transfusion volume will probably bring a postponed surgery for lack of inventory and increase the cost of crossmatch, while the underestimate of transfusion volume will threaten the lives of patients because of the insufficient RBCs supply during surgery. For perioperative patients of surgical patients, an accurate PBV of RBCs plays a particularly important role both from the perspective of patients and hospitals.

The clinicians’ concern is the volume of allogeneic RBCs which are prepared by serological crossmatch prior to the surgery. Reducing unnecessary RBCs transfusion is a big trend. On the one hand, the blood supply source is insufficient [4]. On the other hand, to reduce the incidence of surgical infection and complications, it is common practice to minimize unnecessary blood transfusion. Accumulating evidence indicates that restrictive transfusion at a lower threshold of Hb concentration is safer than liberal transfusion at a higher Hb threshold [5-10]. However, restrictive transfusion could bring some adverse outcomes. For example, the mortality ratio of restrictive to liberal transfusion was 1.64 in a randomized trial [11]. The latest authoritative guidelines from the American Association of Blood Banks (AABB) showed two recommendations for hospitalized adult patients who are hemodynamically stable and neonates based on Hb level [12]. But this guideline is suitable for transfusion recommendation for hemodynamically stable patients, not for pre-surgery blood preparation of surgical patients. To get more accurate PBV of RBCs, the researches on the related pre-surgery indexes of RBC transfusion have made some progresses based on the idea of modeling. Ekhaguere et al.

conducted a retrospective study of very low birth infants over 14 years, and found 13 factors affecting the RBCs transfusion [13]. However, it does not provide specific volume for the decision of clinical transfusion. Liao et. al have given quantitative scoring indicators from a clinical perspective to give the transfusion decision-making of RBCs to make the transfusion more precise [14]. But their model still in the validation phase and the blood transfusion indicators need to be further refined. Fernandes et al. performed a retrospective study on liver transplantation patients to identify pre-surgery indicators affecting blood transfusion [15], but this could not be used as a general model in clinical applications.

The development of artificial intelligence (AI) technology has brought dawn to the medical research. Machine learning is a branch of the field of AI, which application domain is extremely extensive, including almost all human cognitive fields, such as molecular biology, text processing, computer vision and robotics [16]. According to whether the model has the comprehensibility, machine learning models are divided into linear model and nonlinear model. The common linear ones include multivariate linear regression (MLR), and logistic regression (LR) [17]. The common nonlinear ones include support vector machine (SVM) [18], random forest (RF) [19], back propagation neural network (BP) [20], K nearest neighbors (KNN) [21], and XGBoost [22]. These models have been used to predict the pre-surgery transfusion volume or other predictive studies, for example, the random forest algorithm was applied to build the transfusion risk prediction model for the patients undergoing percutaneous coronary intervention [23] and regression model was applied to find the risks of the RBCs' transfusion in obstetric patients [24]. However, the current trained models are for special types of patients or not suitable for predicting the required PBV of the RBCs in surgical patients.

In that this study, we extracted the factors directly related to RBCs transfusion from a large scale of clinical transfusion data. Based on these factors, we learned a more universal model that can be used to predict the required PBV of RBCs in surgical patients.

2 Materials and Methods

2.1 General preparation

The data were collected from electric medical records of patients who underwent surgery of the Chinese PLA General Hospital from January 1, 2011 to December 31, 2016. These records were exported from multiple real-time interaction database systems from the hospital, including HIS, LIS, PACS et al. The medical records included patients' basic demographic information (such as admission number, name, gender, date of birth et al.), laboratory examination (including routine blood work, blood gas, biochemical indicators, blood coagulation function and routine urine analysis), blood transfusion record (such as the application volume of RBCs, the actual transfusion volume of RBCs, the storage of the RBCs et al., which covers all applications and actual transfusion information) and surgery information (such as basic surgical information, pre-surgery medication information, basic disease information et al.).

2.2 Data screening

Quality control for clinical data is a critical step before modeling. It determines the accuracy and reliability of subsequent results in data analysis. Moreover, it's main goal is to identify and correct errors in medical records [25]. The review process includes three repeated cycles of screening, diagnosis, and editing [26]. Based on this idea, a detailed data screening was performed. Criteria for inclusion are below:

1. Only patients who underwent surgery were included.
2. Patients without transfusion or receiving allogenic RBCs were included, and the total RBCs transfusion volume was considered as 0 U for the former, where U represents unit(s), 1unit RBCs is derived from 200 ml whole blood, and 1.5 units RBCs means 300ml blood in China. These patients had already applied for RBCs or submitted application forms.
3. All transfusions with RBCs which occurred during the intra-surgery period were included.
4. Non-duplicate records were included.
5. Missing values were considered unavailable. The entire row of data was retained except that all values of the row were

unavailable. That was, at least one record with a value was retained.

6. The latest laboratory examinations of pre-surgery within 7 days were included.

7. The same patient's RBCs transfusion volumes were merged together when multiple transfusions were within 24 hours in one surgery.

8. Patients whose total transfusion volume exceeded 10U were excluded.

The data tables were extracted and merged manually from raw discrete data tables, where 70,948 surgical no-transfusion and 92,057 surgical RBCs transfusion records were collected, respectively. Finally, 118,823 clean data points with a total of 47,875 transfusions, 39 factors directly related to RBCs transfusion, the actual volume in patients' and clinicians' empirical estimates were extracted. In the subsequent analysis, the data set was divided into two groups, including patients without transfusion and receiving allogenic RBCs. The transfusion volume of RBCs varies from 1U and 10U with an interval of 0.5, where the volumes are discrete and can be divided into 20 categories.

2.3 Machine learning methods

Six typical and common machine learning methods were used to establish the prediction models between the required PBV of RBCs and the data features which directly related to the RBCs transfusion. The methods are LR, MLR, SVM, KNN, BP, RF and XGBoost, where LR is used to build a binary linear classification model, MLR is used to build a multi-linear regression model and the others, including SVM, KNN, BP, RF and XGBoost, are used to build nonlinear models. Their principles are briefly described below. The basic idea of SVM is to find a division hyperplane which separate samples from different categories in the sample space based on training set. SVM belongs to the discriminant model and is based on the efficient sequential minimal optimization algorithm used to solve quadratic programming problems. At present, RF has been widely used because it can compute the variable's nonlinear effect, reflect the interaction between the variables and be not sensitive to outliers. BP is the most successful neural network learning algorithm so far. KNN is also a discriminant learning model, which is widely used because of its simple working mechanism. XGBoost is a relatively new algorithm, which is a cluster of algorithms that can elevate the weak learner to a

strong learner and reduce overfitting effectively.

2.4 Build models

For all models, 90% of data was taken as the training set ($n=106,940$), and the remaining 10% ($n=11,883$) as the test set, where the number of features was 39. To avoid the influence of imbalanced data division in the training model, the training set was generated randomly. Binary classifier means whether a patient have received RBCs (false and true), which corresponds to the decision-making of the pre-surgery preparation of blood. Multi-class classifier means the volume would be considered as classification variable, while linear regression as continuous variable. Multi-class classifier corresponds to the required PBV of RBCs for a patient. Each accuracy was derived from the maximum of the method during adjusting the model's parameters (Table S1). Regarding logistic regression model were built and tested two models, including all the features and only the significant features as input. Then the two models were compared by analysis of variance (ANOVA) (Chi-squared test).

2.5 Model evaluation and validation

In this study, three groups of the volume of RBCs were acquired: the volume the patient applied for, that is clinicians' empirical estimates (experience group), the actual transfusion volume of patients (true group) and the predicted volume generated from the models' outputs (prediction group). The experience group and prediction group were compared to the true group, which would show the performance of the models relative to the clinicians' empirical estimates.

The models were evaluated by the accuracy in classification or regression learning tasks, where the accuracy was calculated by the following formula[17]: $ACC(f; D) = \frac{1}{m} \sum_{i=1}^m \mathbb{I}(f(x_i) = y_i)$. Receiver Operating Characteristic (ROC) plots were used to measure the performance of the machine learning models, where x = False Positive Rate (FPR) and y = True Positive Rate (TPR). When the ROC of a model was completely covered by another model, the latter was better. But when the two ROCs crossed, the Area Under ROC Curve (AUC) was used to measure the performance of a model. To measure linear regression, it was considered acceptable when the error was not greater than 1U according to

the clinicians' experience. The relative importance of the 39 factors were sorted according to the value of MeanDecreaseAccuracy, generated from the random forest model [19].

2.6 Statistical analysis

The correlation between two variables was calculated with Pearson's correlation by IBM Statistics 22. The data report included calculations of the variables' percentage, mean, and interquartile change. T-test and chi-square test were used to test the group difference of two groups between continuous variables and classification variables, respectively. Kruskal-Willis rank sum test was used to test the group difference of more than two groups. Psych package [27] was used for the statistical analysis. The true volume of a patient received was based on the clinicians' experience, and the training of the prediction value was based on the true volume, so the dependent t-test was used to test the group difference of two groups, such as empirical group vs. true group and predicted group vs. true group.

3 Results and Discussion

3.1 Patient characteristics

The detailed demographic and clinical data are shown in Table 1. After the data screening process, a total of 118,823 valid RBCs transfusion records were included in this analysis, with 47,875 (40.3%) having received RBCs. A total of 165,759.5U of RBCs were transfused in 47,875 occurrences (mean, 3.46U). All factors were significantly different between subjects with and without RBCs transfusion except blood group (Table 1). In addition, the significance in different transfusion volume was also tested and the results show that 20 groups who have received RBCs are significant (p -value < 0.001) between 39 factors and the transfusion volume of RBCs.

3.2 Hemoglobin concentration

The distribution of RBCs transfusion was calculated in this study. The frequency of patients who received transfusion was calculated according to pre-surgery Hb concentration, in 6 intervals: (1) $<60\text{g/L}$ (Hb concentration $< 60\text{g/L}$); (2) 60g/L ($60\text{g/L} \leq \text{Hb concentration} < 70\text{g/L}$); (3) 70g/L ($70\text{g/L} \leq \text{Hb concentration} < 80\text{g/L}$); (4) 80g/L ($80\text{g/L} \leq \text{Hb concentration} < 90\text{g/L}$); (5) 90g/L ($90\text{g/L} \leq \text{Hb concentration} \leq 100\text{g/L}$); (6) $>100\text{g/L}$ (Hb concentration $>100\text{g/L}$). The distribution of transfusion rate in different Hb concentration intervals of pre-transfusion is shown Fig. 1. The percentages of the above 6 intervals are 0.7%, 2.8%, 7.8%, 13.7%, 12.0%, 63.0% in pre-transfusion (Fig. S1) and 0.4%, 1.0%, 2.9%,

7.6%, 16.2%, 71.6% in post-transfusion (Fig. S2).

3.3 Factors analysis

The heatmap of the correlation matrix between the 40 factors (include transfusion volume) is shown in Fig. S3. The heatmap shows the degree of relevance between features (39 clinical indicators), where 37 features are significantly correlated with one or more other features except for urine leukocyte (LEU) and urine erythrocyte (ERY) (Pearson correlation when the confidence was 0.05). We also analyzed the effect of storage of RBCs, and the result showed that the storage time of RBCs had a significant positive influence on the volume of RBCs transfusion (Pearson two-tailed test, correlation coefficients: 0.071, p-value < 0.01).

3.4 Modeling and results

Data from this study, were used to test the performance of models generated from 6 typical machine learning algorithms. Two methods of predicting of each machine learning algorithms were binary classifier and multi-classifier. The comparison of accuracies between clinicians' experience recommendation and machine learning models are shown in Fig. 2. All the machine learning models' performance was superior clinicians' experience recommendation (experience group) both in binary and multi-class classifier. The results of two LR models including all the features and only the significant features as input showed that the fitting degree of the two models was the same under different predictive variables ($P=0.2888>0.05$), so the prediction can be preformed with only the significant features as input (RBC, PCO₂, SatO₂, CKI, K⁺, Ca²⁺, INR and LRU were excluded because their regression coefficients were not significant). Regarding MLR model, 4 features including GLU, CKI, ALT and INR were excluded because they would reduce the model's quality in the process of model optimization.

The experience group and prediction group were compared to the true group. Using the dependent T-test, the result showed that True vs. Experience, 95% CI: [1.7807, 1.8818], p-value < 0.001, the mean of the differences is 1.8313; True

vs. Prediction, 95% CI: [-0.8960, -0.8296], p-value < 0.001, the mean of the differences is -0.8628 (Fig. 3). The result showed that the mean of prediction group was much closer to the true group.

According to the results of Fig. 2, RF model obtained the highest accuracy in predicting the required PBV of RBCs under the condition of multi-class classifier. Based on this optimal model, the individual categories' evaluation was performed and features' importance were ranked. The ROC curve of the Random Forest model is shown in Fig. 4a. The accuracy of each class changed from 82.31% to 99.97% (Fig. 4b), which indicated that the accuracy of each category was more than 82%. The indicators considered clinically important are included in the top 20 factors for the RBCs transfusion, like age, OL, weight, Hct, Hb, RBC and height, LYM, BSA and HOD (Fig. 4c).

3.5 Implementation and availability

We developed a lightweight prediction tool, named as PTRBC (Prediction of Transfusion with Red Blood Cells). PTRBC was implemented in Perl and R programming. It is freely available for non-commercial use at <https://github.com/niu-lab/PTRBC>. We also used the R packages from the Comprehensive R Archive Network (<https://cran.r-project.org/>) including DMwR, randmoForest, xgboost, e1071, gmodels and class.

3.6 Discussion

Clinical data mining is an important trend in the development of modern medicine in the information age. The common machine learning solutions include classification, clustering, prediction, and regression [28]. Machine learning is widely used in the field of medicine and has achieved good results, such as the automatic sensing method was developed to sensitively detect unstained malaria-infected RBCs [29]. Machine learning acquires the clinical experience through the means of calculating, although the experience is not entirely correct, the prediction will tend to be more accurate as the data is constantly updated and the model is repeatedly trained. Therefore, to build data-driven decision-making models can help to improve the quality of medical. Before the implementation of the electronic crossmatch, the accurate prediction of the pre-surgical preparation of blood can be realized via calculation, which will effectively save blood

consumption, improve the accuracy of blood supply and relieve the insufficient supply of blood challenge.

In clinical blood transfusion, surgeons often focus on two issues: whether to give the RBCs transfusion and how many a patient should be given. The formula for calculation of blood transfusion volume based on weight and Hb concentration had been extensively studied, specially, for the critically ill patients [30] and children [31]. But these models are for specific groups of patients. Based on the above issues, we have built six models based on the overall clinical context and alternative therapies besides the Hb concentration to predict the transfusion decision and transfusion volume. Our study showed that there were 37 significant factors between the two groups as shown in the last column of Table 1. We found that the rate blood transfusion with RBCs decreased with the increase of hemoglobin concentration (Fig. 1). For patients with transfusion, more than 63.0% of patients received a RBCs transfusion in our cohort, although their Hb concentration of pre-transfusion exceeded 100 g/L (Fig. S1). However, the individualized blood transfusion should be considered for these patients whose Hb concentration of pre-transfusion do not meet the guidelines. The percentage of post-transfusion Hb concentration more than 9g/L is 87.9% (Fig. S2), which means that most patients achieved the recommended Hb concentration in AABB guidelines (9 g/L).

The factors affecting blood transfusion are more complicated, including the patient's own characteristics and the characteristic of the RBCs to be transfused. Specifically, the patient's own factors include multiple blood transfusions, immune system diseases, fatigue, dyspnea, mechanical ventilation, low venous oxygen saturation (SvO_2), blood flow loss, Hb, Hct, inadequate oxygen supply, blood pressure, heart rate, body temperature, and so on [32]. The storage time of blood is also a significant factor [33], and our results show that storage time is also a significant factor. But the storage time of RBCs was not used when building models because our cohort included the patients given more than one bagged RBCs in different storage time in a blood transfusion. In the current report, 39 factors related to RBCs transfusion have been analyzed, and the correlations showed that 37 factors are significant (p -value < 0.05) (Table 1). But in the process of modeling, all 39 factors were all considered in our models. Because the accuracy would decrease with the decrease of the

number of features for the nonlinear machine learning models. Therefore, this study took all features as input for nonlinear machine learning models.

The results showed that random forest and XGBoost models were optimal in predicting the volume of RBCs transfusion in all the training models. The XGBoost model can be used to the assisted decision-making in pre-surgical preparation volume of RBCs (reached 83.1% accuracy), and the random forest model can be used to given the preparation volume of RBCs (reached 72.9% accuracy) (Fig. 2). Nevertheless, the volume of clinicians' empirical estimates only up to 42.9% and 30.4% accuracy, respectively. We observed that the mean value in prediction group was closer to the real group than that in the empirical group (-0.8628 vs. 1.8313) (Fig. 3). The mean value of prediction group was lower than the real group, because there were too many samples who were no-transfusion in the training set, which may lead to some model deviation. Overall, the prediction group had minor fluctuation. Previous studies have been based on Hb and weight to calculate the expected volume of RBCs [30] and the top 10 important factors (Fig. 4c) show that these two indicators are high on the list, where the important factors are consistent with the previous results.

Our model has achieved good results, but there are some minor limitations. Clinically speaking, the model in this study was not applicable for patients who underwent more than 10U RBCs transfusion to minimize the introduce of disturbance. From a modeling perspective, hospital days, though it is a post-surgical factor, was included in our model because it had a high weight on the model.

4 Conclusions

Existing models associated with the required PBV predicting have many shortcomings: not suitable for prediction of pre-surgical RBCs preparation, without generality, and fewer factors were considered. Meanwhile, clinicians experience that used in everyday practice result in lower accuracy, compared with the actual transfusion volume of patients. In this paper, we presented a model combination method and this method can achieve higher accuracy than clinicians experience, not only in pre-surgical blood requirement decision but also in the required volume of allogeneic RBCs. Moreover, a new

prediction tool, PTRBC, is developed in this paper, and this tool can be used to give the specific required PBV of RBCs.

We believe this to be a prospective study predicting the required PBV in surgical patients for the countries who faced with the dilemma of electronic crossmatch, which will provide a novel method for studying other blood components, such as plasma or platelet.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (Grant No. 31771466), the National Key R&D Program of China (Grant nos. 2016YFC0503607, 2018YFB0203903), the Special Project of Informatization of Chinese Academy of Sciences of China (Grant No. XXH13504-08), the Strategic Pilot Science and Technology Project of Chinese Academy of Sciences (Grant No. XDA12010000), the Key Project-subtopic of "13th five-year plan" Military Logistics Research (Grant No. BWS16J006), the Medical Big Data R&D Project of PLA General Hospital (Grant No. 2016MBD-023).

Authors' Contribution This study was designed by D. W., Y. Y. and B. N. R. L. performed the research and drafted the manuscript. X. H. drew all the figures. L. S, Y. F., X. S., Z. D. and X. W. provided the methods and clinical guidance. H. Z. and Y. Z. contributed the code debugging. W. L., X. Y. H., C. D, Z. H, S. C., and X. C. edited and revised the manuscript.

Compliance with Ethical Standards The study protocol was approved by the Ethics Committee of Chinese PLA General Hospital, in accordance with the Declaration of Helsinki.

Conflicts of interest The authors have declared no competing.

References

1. Zhu Y, Xie D, Wang X, Qian K (2017) Challenges and Research in Managing Blood Supply in China. *Transfusion Medicine Reviews* 31 (2):84-88. doi:10.1016/j.tmr.2016.12.002
2. Friedman BA, Burns TL, Schork MA (1980) An analysis of blood transfusion of surgical patients by sex: a question for the transfusion trigger. *Transfusion* 20 (2):179-188. doi:10.1046/j.1537-2995.1980.20280169958.x
3. Klein HG, Spahn DR, Carson JL (2007) Red blood cell transfusion in clinical practice. *Lancet* 370 (9585):415. doi:10.1016/S0140-6736(07)61197-0
4. Zhou WB (2017) Analysis of Causes and Strategies of Unbalanced Supply and Demand of Blood Resource. *China Health Industry*
5. Hébert PC, Wells G, Blajchman MA, Marshall J, Martin C, Pagliarello G, Tweeddale M, Schweitzer I, Yetisir E (1999) A multicenter, randomized, controlled clinical trial of transfusion requirements in critical care. *Transfusion Requirements in Critical Care Investigators, Canadian Critical Care Trials Group. New England Journal of Medicine*

- 340 (6):409-417. doi:10.1056/NEJM199902113400601
6. Hill S, Carless PA, Henry DA, Carson JL, Hebert PP, Henderson KM, Mcclelland B (2000) Transfusion thresholds and other strategies for guiding allogeneic red blood cell transfusion. John Wiley & Sons, Ltd. doi:10.1002/14651858.CD002042
 7. Carson JL, Terrin ML, Noveck H, Sanders DW, Chaitman BR, Rhoads GG, Nemo G, Dragert K, Beaupre L, Hildebrand K (2011) Liberal or restrictive transfusion in high-risk patients after hip surgery. *New England Journal of Medicine* 365 (26):2453. doi:10.1056/NEJMoa1012452
 8. de Gast-Bakker DH, de Wilde RB, Hazekamp MG, Sojak V, Zwaginga JJ, Wolterbeek R, De JE, Bj GVDV (2013) Safety and effects of two red blood cell transfusion strategies in pediatric cardiac surgery patients: a randomized controlled trial. *Intensive Care Medicine* 39 (11):2011-2019. doi:10.1007/s00134-013-3085-7
 9. Holst LB, Haase N, Wetterslev J, Wernerman J, Guttormsen AB, Karlsson S, Johansson PI, Aneman A, Vang ML, Winding R (2014) Lower versus higher hemoglobin threshold for transfusion in septic shock. *New England Journal of Medicine* 371 (15):1381. doi:10.1056/NEJMoa1406617
 10. Lee SJ, Seo H, Kim HC, Lim SM, Yoon SJ, Kim HS, Ku JH, Park HP (2015) Effect of Intraoperative Red Blood Cell Transfusion on Postoperative Complications After Open Radical Cystectomy: Old Versus Fresh Stored Blood. *Clinical Genitourinary Cancer* 13 (6):581-587. doi:10.1016/j.clgc.2015.06.002
 11. Reeves BC, Rogers CA, Murphy GJ (2015) Liberal or Restrictive Transfusion after Cardiac Surgery. *New England Journal of Medicine* 373 (2):191. doi:10.1056/NEJMx150020
 12. Carson JL, Guyatt G, Heddle NM, Grossman BJ, Cohn CS, Fung MK, Gernsheimer T, Holcomb JB, Kaplan LJ, Katz LM (2016) Clinical Practice Guidelines From the AABB: Red Blood Cell Transfusion Thresholds and Storage. *Jama* 316 (19):2025. doi:10.1001/jama.2016.9185
 13. Ekhaguere OA, Jr MFH, Bell EF, Nadkarni P, Widness JA (2016) Predictive Factors and Practice Trends in Red Blood Cell Transfusions for Very Low Birth Weight Infants. *Pediatric Research* 79 (5):736-741. doi:10.1038/pr.2016.4
 14. Liao R, Liu J (2014) Index score in Huaxi perioperative blood transfusion - a blood transfusion score based on clinical needs. *Chin J Clin Thorac Cardiovasc Surg* (2):145-146 (In Chinese)
 15. Fernandes DS, Pereira Real CC, Pa SCR, Fb MCDB, Marques Aragão IM, Guimarães Fonseca LF, Gonçalves Aguiar JM, Costa Branco TM, Fernandes Moreira ZM, Barros Esteves SM (2017) Pre-operative predictors of red blood cell transfusion in liver transplantation. *Blood Transfus* 15 (1):1-4. doi:10.2450/2016.0203-15
 16. Robert C (2012) *Machine Learning, a Probabilistic Perspective*. MIT Press. doi:10.1080/09332480.2014.914768

17. Zhou Z (2016) Machine learning. Tsinghua University Press. ISBN:978-7-302-42328-7
18. Dimitriadou E, Hornik K, Leisch F, Meyer D, Weingessel A (2011) Misc functions of the Department of Statistics (e1071), TU Wien
19. Liaw A, Wiener M (2002) Classification and Regression by randomForest. R News 23 (23)
20. Ripley WNVaBD (2002) Modern Applied Statistics with S. Fourth Edition. Springer, New York. ISBN:0-387-95457-0
21. Warnes GR (2015) gmodels: Various R Programming Tools for Model Fitting
22. Chen T, Guestrin C (2016) XGBoost: A Scalable Tree Boosting System. In: ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp 785-794. doi:10.1145/2939672.2939785
23. Gurm HS, Kooiman J, Lalonde T, Grines C, Share D, Seth M (2014) A random forest based risk model for reliable and accurate prediction of receipt of transfusion in patients undergoing percutaneous coronary intervention. Plos One 9 (5):e96385. doi:10.1371/journal.pone.0096385
24. Shehata N, Chassé M, Colas JA, Murphy M, Forster AJ, Malinowski AK, Ducharme R, Fergusson DA, Tinmouth A, Wilson K (2017) Risks and trends of red blood cell transfusion in obstetric patients: a retrospective study of 45,213 deliveries using administrative data. Transfusion 57. doi:10.1111/trf.14184
25. Rahm E, Hong HD (2000) Data Cleaning: Problems and Current Approaches. IEEE Data Engineering Bulletin 23 (23):3-13
26. Van dBJ, Cunningham SA, Eeckels R, Herbst K (2005) Data cleaning: detecting, diagnosing, and editing data abnormalities. Plos Medicine 2 (10):e267. doi:10.1371/journal.pmed.0020267
27. Revelle W (2013) psych: Procedures for Psychological, Psychometric, and Personality Research. R Package Version 1.0-95
28. Bilal M, Oyedele LO, Qadir J, Munir K, Ajayi SO, Akinade OO, Owolabi HA, Alaka HA, Pasha M (2016) Big Data in the construction industry: A review of present status, opportunities, and future trends. Advanced Engineering Informatics 30 (3):500-521. doi:10.1016/j.aei.2016.07.001
29. T G, JH K, H B, SJ L (2018) Machine learning-based in-line holographic sensing of unstained malaria-infected red blood cells. Journal of biophotonics:e201800101. doi:0.1002/jbio.201800101
30. Morris KP, Naqvi N, Davies P, Smith M, Lee PW (2005) A new formula for blood transfusion volume in the critically ill. Archives of Disease in Childhood 90 (7):724-728. doi:10.1136/adc.2004.062174
31. Davies P, Robertson S, Hegde S, Greenwood R, Massey E, Davis P (2007) Calculating the required transfusion volume in children. Transfusion 47 (2):212-216. doi:10.1111/j.1537-2995.2007.01091.x

32. Vincent JL (2012) Indications for blood transfusions: too complex to base on a single number?. *Annals of Internal Medicine* 157 (1):71-72. doi:10.1059/0003-4819-156-12-201206190-00431
33. Koch CG, Li L, Sessler DI, Figueroa P, Hoeltge GA, Mihaljevic T, Blackstone EH (2008) Duration of red-cell storage and complications after cardiac surgery. *New England Journal of Medicine* 358 (12):1229-1239. doi:10.1056/NEJMoa070403

Figure legends

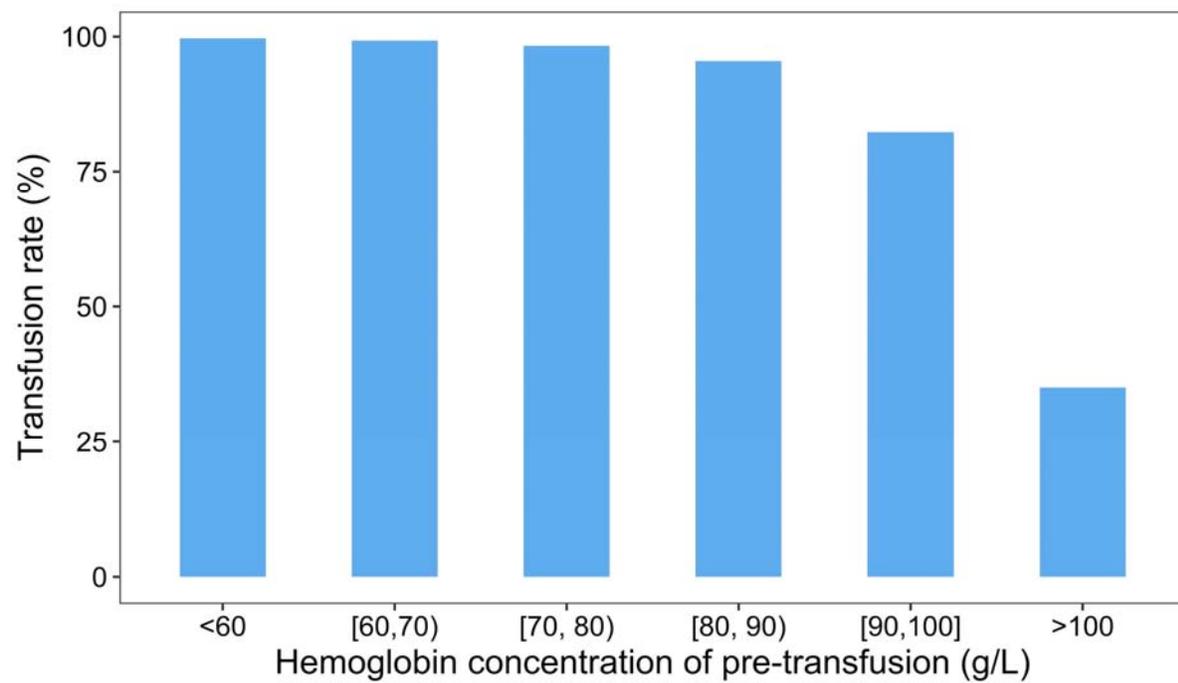


Fig. 1 Transfusion rate in different hemoglobin concentration intervals of pre-transfusion.

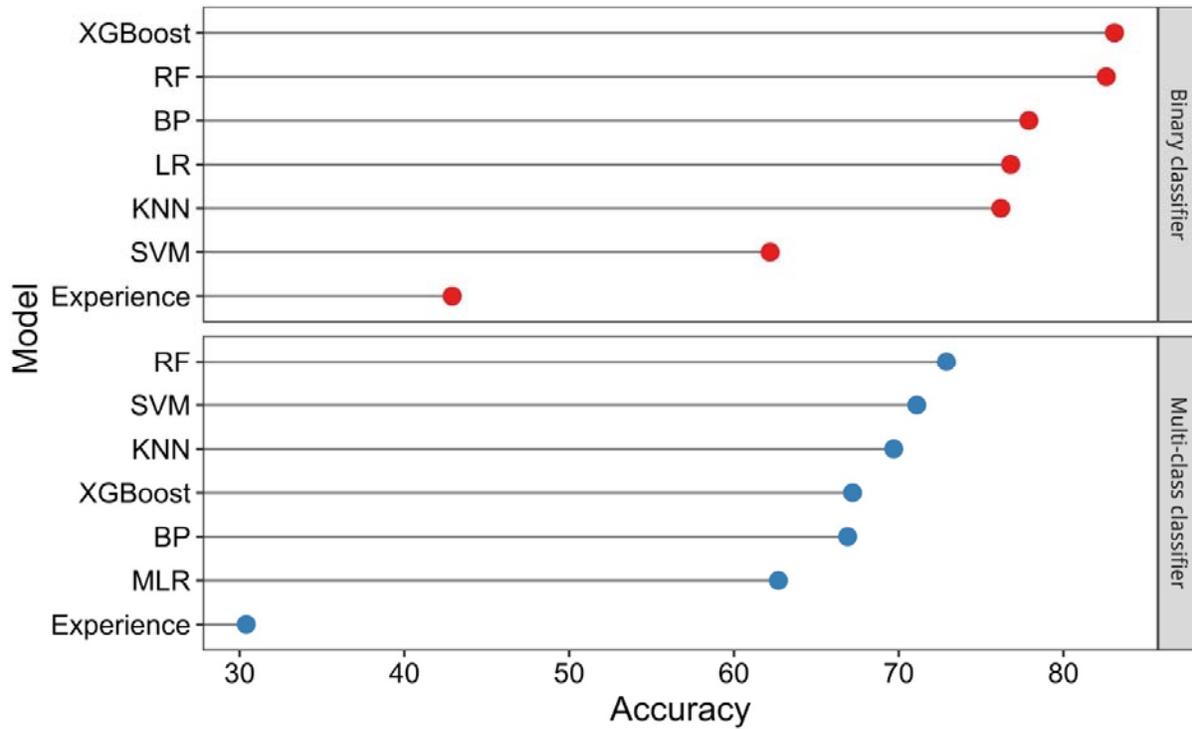


Fig. 2 Comparison of accuracies between clinicians' experience recommendation and models. In multi-class classifier group, the accuracies are normalized values in regression. "Experience" represents the clinician's recommendation.



Fig. 3 Comparison of the difference of the volume RBCs between true vs. experience and prediction.

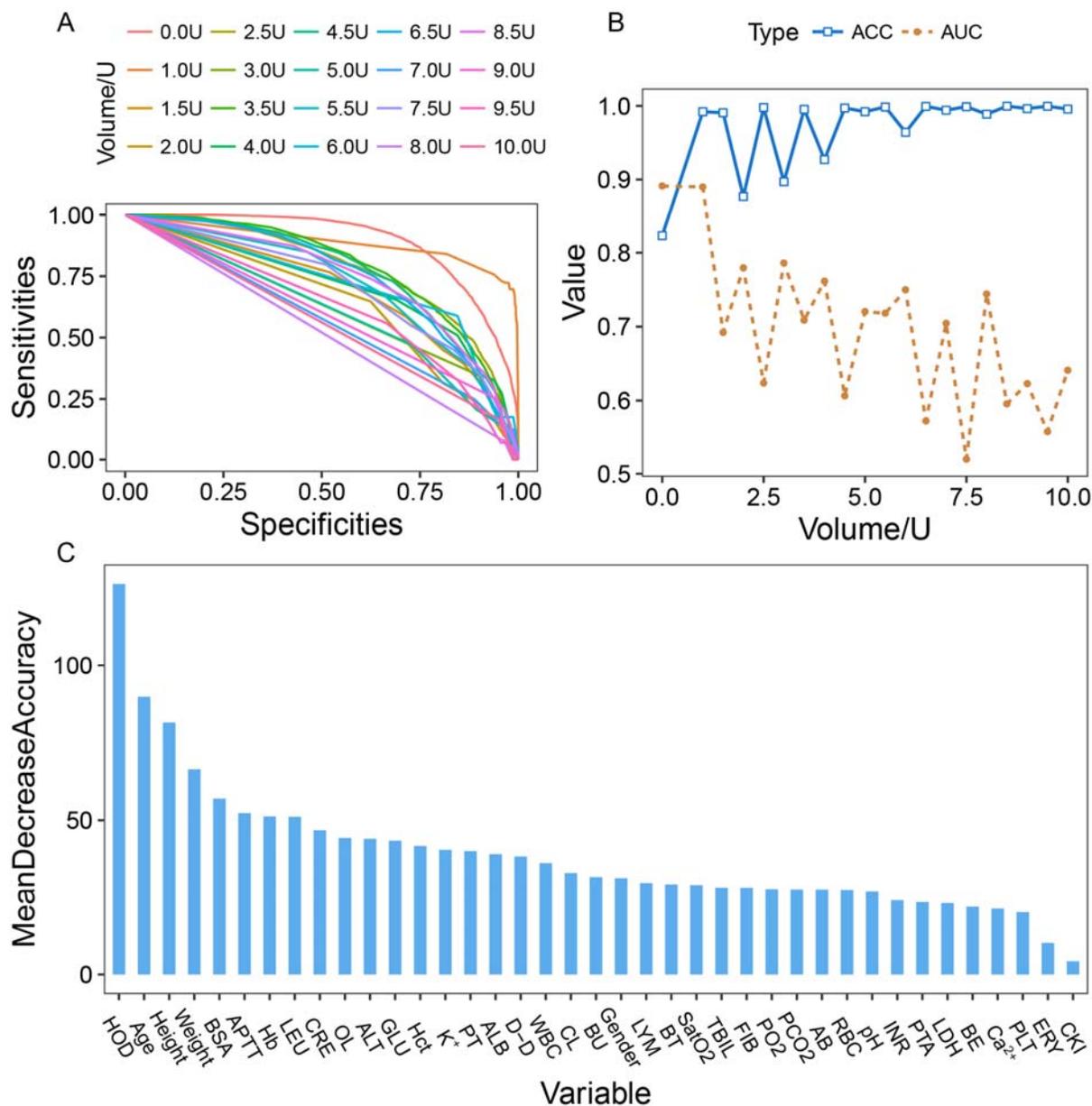


Fig. 4 **a** Receiver Operating Characteristic (ROC) curve of each class in the Random Forest model. **b** Comparison of area

under the curve (AUC) and accuracy (ACC) in each class. **c** Importance of factors in the Random Forest model.

MeanDecreaseAccuracy means the reduction in the degree of accuracy when the value of a variable is changed into a random number. The larger the value is, the greater the importance of the variable.

Table 1 Characteristics and clinical features of the surgical patients

Variable *	All patients (N=118, 823)	Patients without transfusion (N=70, 948)	Patients with transfusion (N=47, 875)	95 percent confidence interval / X-squared, degree of freedom (df)	p-value †
Volume: Mean (1st Qu-3rd Qu) ‡, units	1.395 (0.0-3.0)	0.0 (0.0-0.0)	3.463 (2.0-4.0)		
Age: Mean (1st Qu-3rd Qu), years	50.7 (40.0-63.0)	49.55 (39.00-61.00)	52.4 (42.0-66.0)	[-3.0623, -2.6531]	< 0.0001 §
Sex: no. (%)				293.9600, df = 1	< 0.0001
Male	57957 (48.78)	36055 (50.82)	21902 (45.75)		
Female	60866 (51.22)	34893 (49.18)	25973 (54.25)		
Height: Mean (1st Qu-3rd Qu), cm	164.0 (160.0-171.0)	165.3 (160.0-172.0)	161.9 (158.0-170.0)	[3.2243, 3.5428]	< 0.0001 §
Weight: Mean (1st Qu-3rd Qu), kg	66.36 (58.00-75.00)	67.8 (60.0-76.0)	64.22 (56.00-74.00)	[3.4013, 3.7511]	< 0.0001 §
Body surface area (BSA): Mean (1st Qu-3rd Qu), m2	2.0332 (1.9042-2.1917)	2.0529 (1.9178-2.2003)	2.0040 (1.8836-2.1803)	[0.0456, 0.0521]	< 0.0001 §
Blood type (BT): no. (%)				6.8551, df = 3	0.0767
No laboratory results ¶	5811 (4.89)	4160 (5.86)	1651 (3.45)		
A	30832 (25.95)	18162 (25.60)	12670 (26.47)		
AB	11443 (9.63)	6886 (9.71)	4557 (9.51)		
B	36912 (31.06)	21833 (30.77)	15079 (31.50)		
O	33825 (28.47)	19907 (28.06)	13918 (29.07)		
Hospital days (HOD): Mean (1st Qu-3rd Qu), days	16.28 (10.00-19.00)	14.11 (9.00-18.00)	19.5 (11.0-22.0)	[-5.5562, -5.2199]	< 0.0001 §
Operation level (OL): no. (%)				1115.3000, df = 3	< 0.0001
No laboratory results ¶	63 (0.05%)	31 (0.04%)	25 (0.05%)		
Super-large	28579 (24.05%)	14751 (20.79%)	13828 (28.89%)		
Major	78976 (66.47%)	49503 (69.77%)	29473 (61.57%)		
Secondary	9692 (8.16%)	5642 (7.95%)	4050 (8.46%)		
Minor	1513 (1.27%)	1021 (1.45%)	492(1.03%)		
White blood cell count (WBC): Mean (1st Qu-3rd Qu), x10 ⁹ /L	6.932 (5.090-7.850)	6.294(4.990-7.220)	7.699(5.270-8.900)	[-1.4485, -1.3608]	< 0.0001 §
Platelet count (PLT): Mean (1st Qu-3rd Qu), x10 ⁹ /L	219.8 (170.0-262.0)	222.2 (179.0-259.0)	217.0 (156.0-266.0)	[4.1097, 6.2743]	< 0.0001 §
Leukomonocyte (LYM): Mean (1st Qu-3rd Qu), %	0.285 (0.207-0.367)	0.321 (0.262-0.381)	0.243 (0.133-0.338)	[0.0768, 0.0798]	< 0.0001 §
Red blood cell count (RBC): Mean (1st Qu-3rd Qu), x10 ¹² /L	4.152 (3.780-4.660)	4.446 (4.110-4.780)	3.80(3.11-4.43)	[0.6360, 0.6546]	< 0.0001 §

* All indicators were measured before transfusion;

† For all analyses, p < 0.05 was considered significant; The significance was the difference between patients without and with transfusion;

‡ 1st Qu-3rd Qu means interquartile change;

§ T-test was applied;

|| Chi-squared test was applied;

¶ Unavailability.

Table 1 Continued.

Variable	All patients (N=118, 823)	Patients without transfusion (N=70, 948)	Patients with transfusion (N=47, 875)	95 percent confidence interval / X-squared, degree of freedom (df)	p-value
Hemoglobin (Hb): Mean (1st Qu-3rd Qu), g/L	124.2(110.0-141.0)	134.5(124.0-146.0)	111.9(89.0-132.0)	[22.3574, 22.9409]	< 0.0001 [§]
Hematocrit (Hct): Mean (1st Qu-3rd Qu), L/L	0.370 (0.334-0.417)	0.399 (0.370-0.430)	0.336 (0.275-0.393)	[0.06184, 0.0635]	< 0.0001 [§]
pH: Mean (1st Qu-3rd Qu)	7.36 (7.34-7.38)	7.36 (7.34-7.38)	7.37 (7.34-7.39)	[-0.0077, -0.0061]	< 0.0001 [§]
Partial pressure of carbon dioxide (PCO ₂): Mean (1st Qu-3rd Qu)	43.76 (40.80-47.20)	44.38 (41.60-47.60)	42.93 (39.80-46.50)	[1.3257, 1.5760]	< 0.0001 [§]
Residue determination (BE): Mean (1st Qu-3rd Qu)	1.83 (0.70-2.50)	1.73 (0.60-2.40)	1.96 (0.70-2.60)	[-0.2748, -0.1884]	< 0.0001 [§]
Determination of oxygen partial pressure (PO ₂): Mean (1st Qu-3rdQu)	89.28 (78.80-95.80)	89.8 (79.9-96.2)	88.56 (77.30-95.00)	[1.3257, 1.5760]	< 0.0001 [§]
Actual bicarbonate radical (AB): Mean (1st Qu-3rdQu)	24.29 (23.10-25.70)	24.47 (23.30-25.80)	24.05 (22.80-25.60)	[0.3460, 0.4752]	< 0.0001 [§]
Oxygen saturation measurement (SatO ₂): Mean (1st Qu-3rd Qu)	96.17 (95.60-97.30)	96.27 (95.70-97.30)	96.04 (95.40-97.30)	[0.1697, 0.2967]	< 0.0001 [§]
Albumin (ALB): Mean (1st Qu-3rd Qu)	40.09 (37.90-43.20)	41.31 (39.20-43.60)	37.17 (32.90-41.60)	[4.0484, 4.2286]	< 0.0001 [§]
Serum glucose (GLU): Mean (1st Qu-3rd Qu)	5.42 (4.50-5.60)	5.173 (4.480-5.360)	6.034 (4.600-6.580)	[-0.8975, -0.8226]	< 0.0001 [§]
Serum creatine kinase isoenzyme (CKI): Mean (1st Qu-3rd Qu)	15.8 (11.5-17.3)	15.55 (11.50-17.20)	16.67 (11.40-18.10)	[-1.5783, -0.6634]	< 0.0001 [§]
Serum potassium (K ⁺): Mean (1st Qu-3rd Qu)	4.02 (3.78-4.23)	4.017 (3.790-4.210)	4.025 (3.750-4.260)	[-0.0142, -0.00179]	0.0116 [§]
Total bilirubin (TBIL): Mean (1st Qu-3rd Qu)	17.46 (7.70-14.30)	16.11 (7.90-14.00)	20.79 (7.30-15.70)	[-5.3970, -3.9574]	< 0.0001 [§]
Creatinine (CRE): Mean (1st Qu-3rd Qu)	71.99 (56.50-79.40)	71.11 (58.30-80.10)	74.08 (51.90-76.80)	[-4.5515, -1.3828]	0.0002 [§]
Alanine aminotransferase (ALT): Mean (1st Qu-3rd Qu)	29.7 (11.6-27.3)	26.63 (11.80-26.50)	37.01 (11.10-29.80)	[-12.0250, -8.7386]	< 0.0001 [§]
Lactic dehydrogenase (LDH): Mean (1st Qu-3rd Qu)	167.9 (132.6-174.6)	155.4 (131.6-167.9)	198.3 (136.4-198.9)	[-47.10688, -38.69385]	< 0.0001 [§]
Chloride (CL): Mean (1st Qu-3rd Qu)	104.2 (102.2-106.3)	104.1 (102.3-106.1)	104.3 (102.0-106.8)	[-0.2932, -0.1680]	< 0.0001 [§]
Blood calcium (Ca ²⁺): Mean (1st Qu-3rd Qu)	2.24 (2.17-2.32)	2.261 (2.190-2.330)	2.194 (2.090-2.300)	[-0.2932, -0.1680]	< 0.0001 [§]
Serum urea (BU): Mean (1st Qu-3rd Qu)	5.25 (3.96-5.97)	5.043 (4.000-5.850)	5.755 (3.870-6.370)	[-0.7702, -0.6537]	< 0.0001 [§]
Prothrombin time (PT): Mean (1st Qu-3rd Qu)	13.42 (12.60-13.80)	13.12 (12.50-13.60)	13.93 (12.70-14.40)	[-0.8445, -0.7756]	< 0.0001 [§]

Table 1 Continued.

Variable	All patients (N=118, 823)	Patients transfusion (N=70, 948)	without Patients transfusion (N=47, 875)	with 95 percent confidence interval / X-squared, p-value degree of freedom (df)	
Activated partial thromboplastin time (APTT): Mean (1st Qu-3rd Qu)	37.9 (34.2-40.0)	36.97 (34.00-39.30)	39.53 (34.50-41.60)	[-2.6891, -2.4273]	< 0.0001 [§]
Prothrombin activity (PTA): Mean (1st Qu-3rd Qu)	93.82 (85.00-104.00)	96.64 (89.00-105.00)	88.68 (78.00-100.00)	[7.5667, 8.3627]	< 0.0001 [§]
International normalized ratio (INR): Mean (1st Qu-3rd Qu)	1.06 (0.98-1.09)	1.03 (0.97-1.07)	1.105 (0.990-1.150)	[-0.0815, -0.0734]	< 0.0001 [§]
Fibrinogen (FIB): Mean (1st Qu-3rd Qu)	3.41 (2.65-3.91)	3.28 (2.61-3.70)	3.67 (2.77-4.34)	[-0.4194, -0.3598]	< 0.0001 [§]
D-Dimer (D-D): Mean (1st Qu-3rd Qu)	1.61 (0.32-1.52)	0.73 (0.28-0.61)	2.70 (0.43-3.29)	[-2.0647, -1.8599]	< 0.0001 [§]
Urine leukocyte (LEU): Mean (1st Qu-3rd Qu)	61.47 (2.40-19.80)	53.84 (2.20-17.40)	78.21 (2.90-25.90)	[-32.8519, -15.8995]	< 0.0001 [§]
Urine erythrocyte (ERY): Mean (1st Qu-3rd Qu)	124.8 (2.8-11.8)	105.6 (2.7-11.5)	164.7 (3.0-12.4)	[-86.4840, -31.6662]	< 0.0001 [§]