Prediction model for Hemodilution Effect among patients with Fluid Resuscitation

Final Report

Team H

Abstract

Background- The administration of intravenous fluid (resuscitation) is one of the most common interventions in intensive care medicine. Decreased concentration of hemoglobin in intravascular plasma volume due to IV fluid infusion is called hemodilution. Clear guidelines do not exist for determination of the expected amount of hemodilution after fluid administration. The main goal of this study is to qualify and identify any correlation in the change in hemoglobin to the amount of fluid administered, and to evaluate also for any other factors that may contribute to hemodilution.

Methods- Using the MIMIC III database (2001 – 2012), a study cohort of patients from 18 – 80 years of age, who didn't have a diagnosis of internal or external bleeding or any bleeding disorder, and were given normal saline within the first 24 hours of their ICU stay, was formed. Covariates like gender, weight, urine output over 6 hours, creatinine levels at admission, hemoglobin at baseline and after 24 hours, and total amount of fluid administered were collected. Six prediction models were built, viz., linear regression, stepwise linear regression (AIC), random forest, support vector machine, gradient boosting machine, neural network.

Results- Comparing mean square error of all six models, SVM with MSE of 1.02g/dl was finalized as best model. Model performance is better near the median change in hemoglobin among the cohort and has increasing error on both extremes.

Conclusion- Our study has proven that hemodilution exists in the critically ill cohort using normal saline over the first 24 hours, but there is either no simple model to predict the expected drop or we are limited by the data.

Introduction

Hemoglobin is a tetrameric protein for transport of oxygen to organs and tissues and is contained in Red Blood Cells (RBCs). It is an intravascular molecule that, in the absence of bleeding, typically remains within the circulation. Anemia is defined as a reduction in the hemoglobin level in the blood. Major causes of anemia are 1) acute blood loss by internal or external bleeding, 2) deficient RBC production like in leukemia 3) increased RBC destruction or sequestration such as in splenomegaly or sickle cell anemia [1] and 4) dilutional anemia. Decreased concentration of hemoglobin in intravascular plasma volume due to IV fluid infusion is called hemodilution [2]. The administration of intravenous fluid (resuscitation) is one of the

most common interventions in intensive care medicine, directed mainly at improving microcirculation to attain better oxygen delivery [3].

The concept of hemodilution is largely unexplored in humans. Several animal studies show that fluid administration as a treatment for hemorrhage results in physiological hemodilution [4,5,6,7]. In human studies, the focus is mostly on acute normovolemic hemodilution (ANH) in cardiac surgery [17]; the hemodilution effect has been observed in cardiac surgery during cardiopulmonary bypass due to fluids [8]. Others have attempted to create prediction models that calculate the hemodilution rate due to bypass surgery [9]. However, weak correlation between fluid resuscitation and hemoglobin among septic shock patients has been refuted by one study [10]. One study of 251 ICU patients also implies the fluid status would have an effect on the hemoglobin concentration [16]. One prospective study of 84 patients in Shiraz, Iran found that changes in clinical or biochemical parameters like hemoglobin, WBC, platelet count, HCO3, PCO2 were statistically significant after 1 liter of saline administration in patients with minor trauma [11].

Another recent study considered volunteer blood donation a proxy for acute hemorrhage and quantified hemodilution after crystalloid fluid administration and built a prediction model for change in hemoglobin concentration. However, blood donation as a proxy is limited. The amount of trauma by donation is not completely comparable to actual trauma, as the former excludes inflammatory release, systemic response by the body, continuous losses of blood, and input of fluids. Furthermore, blood donors are inherently a healthier, younger population and the predicted association is assumed linear in nature in this study [12].

Accordingly, a clear guideline does not exist for determination of the expected amount of hemodilution after fluid administration. The main goal of this study is to qualify and identify any correlation in the change in hemoglobin to the amount of fluid administered, and to evaluate also for any other factors that may contribute to hemodilution. If an estimation of hemoglobin change can be made after fluid administration, the index could be used to differentiate dilutional anemia from other causes of anemia (e.g. active bleeding) that may need urgent and invasive procedures. We could subsequently detect abnormal changes earlier and provide better treatment to patients.

Methods

Data Sources:

Our study is based on hospital-based, open, prospective data collected in the MIMIC-III (Medical Information Mart for Intensive Care III) database which has information of over 40,000 patients who stayed in critical care units of the Beth Israel Deaconess Medical Center between years 2001 and 2012. The database includes information such as demographics, vital sign measurements made at the bedside (1 data point per hour), laboratory test results, procedures, medications, caregiver notes, imaging reports, and mortality (both in and out of hospital) [13,14]. We used this data to build predictive models and for cross-validation. All the data in MIMIC-III do not contain any patient protected health information (PHI).

Study Cohort:

We obtained data from the MIMIC dataset, including patients whose age was between 18 to 80 and who received only normal saline (without other crystalloid or colloid) during the first 24 hours of their ICU stay. We then excluded patients who were diagnosed with unstable hemodynamic or active bleeding status with ICD-9 codes, given that it may alter the concentration of hemoglobin significantly. Patients who were readmitted to the ICU during a single hospitalization were not included.

[Table 1: Selection of cohort]

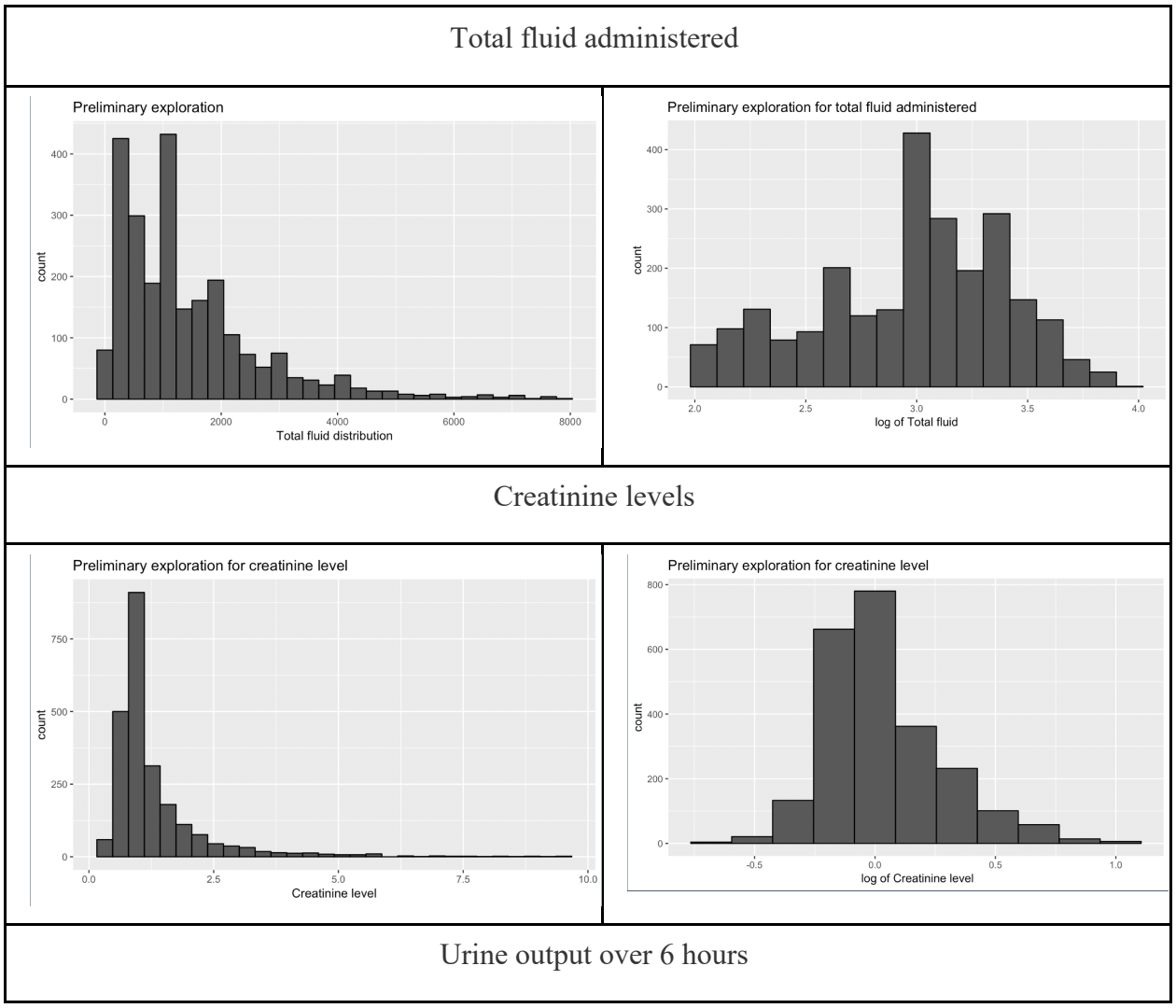
Inclusion and Exclusion Criteria	Number
Age 18 to 80	41091
Receiving NS within 24 hrs	16088
Excluding patients with unstable hemodynamic status or active bleeding	7015
Excluding patients who received crystalloid other than NS, or any colloid	4093
Patients who have Hgb measured within 6 hours before ICU admission	3470
Patients who have Hgb measured within the first 24 hours of ICU stay	3190
Patients without missing data*	2462

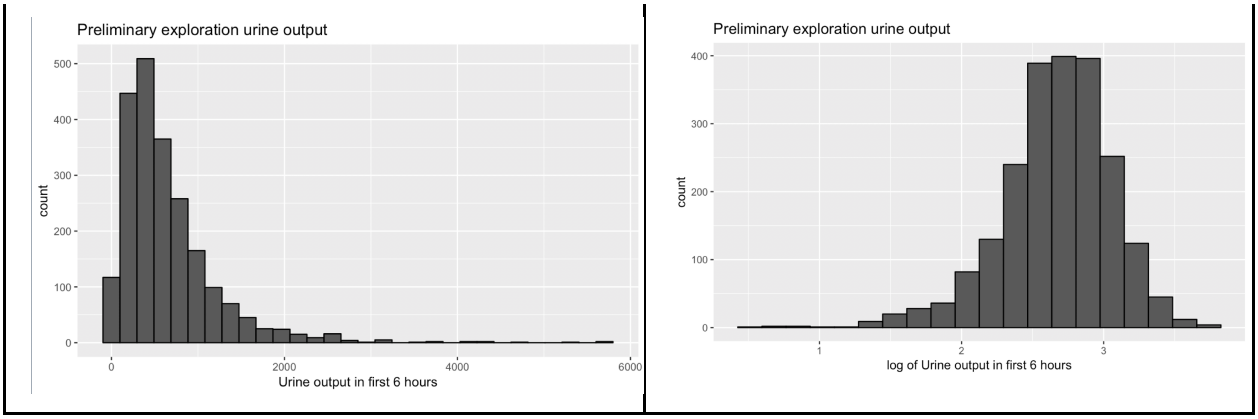
Data*: age, gender, ethnicity,total urine out of 6 hours, creatinine of 1st hour, ventilation setting(mechanical ventilation, oxygen therapy, self-extubation, extubation), the duration between first and last measurements of hemoglobin within 24 hours, the total volume of normal saline within 24 hours.

Our final cohort had 16 variables. We considered the difference between the hemoglobin levels from first to last measurement as an outcome variable. Since hemoglobin was not measured at exactly 24 hours after admission for all patients, we included the time between first and last measurement as a variable. Weight, admission creatinine level and urine output over 6 hours had missing data. We assumed all three variables were missing values at random (MAR). Missing data was imputed by using missRanger package in R [18] which utilizes the chain tree ensembles for fast imputation of missing values.

Distribution of individual covariates were explored. To obtain a better linear model, we log transformed (base 10) total fluid given to the patient, creatinine levels at the first measurement after admission to ICU and urine output measured over the first 6 hours to obtain approximately normal distributions [Figure 1]. Age had a non-normal distribution; however, transforming the age variable did not approximate the distribution to normal. We hence decided to not transform the age variable. We also decided to exclude a total of 8 patients with total fluid administration above 8 liters within first 24 hours to avoid excessive influence by these outliers.

[Figure 1: Comparison of distribution of log transformed variables]





Data Preparation

We divided dataset in two subsets, namely, "training set" with a randomly selected 70% of complete data and "test set" with the remaining 30% of the complete data. To adjust for different scales for covariates, training set was normalized using Min-Max formula [Figure 2]. The minimum and maximum values for each column of the training dataset were then used to normalize the test dataset.

[Figure 2: Min- Max formula for normalization]

$$z = \frac{x - \min(x)}{\max(x) - \min(x)}$$

Modelling:

Six statistical models were compared for both transformed and non-transformed data. All models were trained based on the "training subset" and performance was tested based on predictions for the "test subset". Overall, all models performed better without transformations. [Please find individual results of log transformed models in supplement 1]. Models built were as follows:

1. Model 1 - Generalized linear regression was performed with inclusion of all predictors.

2. Model 2 - Best predictive model was selected based on Stepwise Akaike integration criteria (AIC) method. Final model was selected as one with smallest AIC and the variables in the model were considered as best predictors.

3. Model 3 - Random forest (RF) model. Ensemble technique proposed by Breiman et al in 2001 was used [19]. This methods utilized two important principles namely bagging and random feature selection. Within the "training data set", number of trees to be built (ntrees) were selected and for each tree, a bootstrapped sample from the training set was generated. Using these bootstrapped samples, an unpruned regression tree was built where for each split (mtry) number of variables were selected from all original variables.

The best split point for the variables was selected to maximize the information gain measure by Gini impurity and two child nodes were formed in the tree.

The rest of training data, from out of bucket bootstrapped samples is used to obtain test error and to calculate the importance of each variable. For prediction, input features pass through from the root to the end nodes of all trees based on the predetermined splits. The output of each tree is defined as the average objective value in the end nodes and the ensemble average of outputs from all the trees is considered as the final estimate.

To improve learning of RF algorithm, we can tune several different parameters. To tune our model, we created a hyper-grid with different combinations of number of trees (500), the number of variables to randomly sample as candidates at each split(ranging from 0 to 15), proportion of samples to train on (0.55, 0.632, 0.70, 0.80), minimum number of samples within the terminal nodes (ranging from 2 to 15). A total of 896 combinations were tested on the training set and the model was evaluated on test set. R package "ranger" was used for this analysis [20].

4. Model 4 - Support vector machine model (SVM). When correlation between predictors and dependent outcome variable is linear; support vector regression (SVR) has a similar equation form to multiple linear regression. However, the SVR method has two additional parameters: *C* and ε. The first parameter *C* is introduced to adjust the error sensitivity of the training data in order to avoid over-fitting; setting *C* to a high value results in fewer prediction errors in the training data. The second parameter ε is the regularization constant, which controls the flatness of the final model. *j* represents the *j* th variable [Figure 3].

[Figure 3: Support vector regression equation]

$$\min \sum_{i=1}^{m} (y_i - (\hat{W} \cdot X_i + \hat{b}))^2 + C \sum_{j=1}^{n} |W_j^2|,$$

The goal of SVR is to determine an optimal function, that has a regression hyperplane as flat as possible, and also has less than ε deviation from the target values for the training data, such that we do not count errors that are less than ε. By using different kernel functions, which transform data into a high dimensional space or add non-linearity, the SVR algorithm allows application of nonlinear regression.

In this study we did extensive grid search with combinations of ε (ranging from 0 to 1 with jumps of 0.1) and C (sequence of whole numbers of 1 to 100). Radial kernel with gamma values of 0.0005, 0.0001, 0.005, 0.05, 0.5 were included. The parameters with the best model performance were selected. By using different kernel functions, which transform data into a high dimensional space or add non-linearity, the SVR algorithm allows application of nonlinear regression. R package "e1071" was used [21].

5. Model 5 - Gradient Boosting machine model (GBM). The main idea of boosting is to sequentially add new models to the ensemble. At each particular iteration, a new weak, base-learner model is trained with respect to the error of the whole ensemble learnt so far. The model fits a first decision tree followed by a second tree fit for the residuals of the first tree and is added to the ensemble. This process is then continued. A cross validation mechanism is used to stop the boosting.

   Gradient boosting minimizes the mean squared error (MSE) loss function. Gradient descent function is used to tweak parameters iteratively in order to minimize a cost function. *Stochastic gradient descent* is used for sampling a fraction of the training observations (typically without replacement) and growing the next tree using that subsample. We can tune multiple parameters of this model to find the best predictive model.

   In this study we did a hyper-grid search to find the best model. We searched across 81 models with varying learning rates and tree depths. We used an interaction depth of 1, 3 and 5 nodes and a learning rate of 0.01, 0.1, and 0.3. We also varied the minimum number of observations allowed in the trees terminal nodes and introduced stochastic gradient descent by allowing bagging fraction of 0.65, 0.8 and 1. Looping through each parameter combination, we applied 5000 trees. We used the training set for this purpose and then evaluate the models on test set. After obtaining the best model we used the parameters for the model and cross validated the model. R package “gbm” was used [22].

6. Model 6 - Neural Network model. Artificial Neural Networks (ANN) are composed of an interconnected group of artificial neurons (ANs) which follow an interconnection schema with high density and parallelism, and have weights (synapses) whose modification represents network learning. We used grid search with possibility of 5, 6, 7 hidden layers and decay of 0.5 and 0.1. Model was trained on the training set and then evaluated on test set. R package “caret” was used [23].

Results:

The distribution of all the covariates in our cohort is shown in Table 2. 20.43% of final cohort had missing data for weight, 3.04% of final cohort had missing data for creatinine level and 10.96% of final cohort had missing data for urine output over 6 hours. The distribution of the variables after imputation are shown in Table 3. Total fluid administered to the patient, creatinine levels(mg/dl) at the first measurement after admission to ICU and urine output measured over the first 6 hours had the right tailed distribution and hence were log transformed [Figure 1]. The Pearson’s correlation coefficient between difference in hemoglobin levels and total fluid given

was -0.14 and was statistically significant with p value <0.0001. Mean squared errors (MSE) for all six models were compared and support vector machine model was selected as final model.

[Table 2: Cohort characteristics]

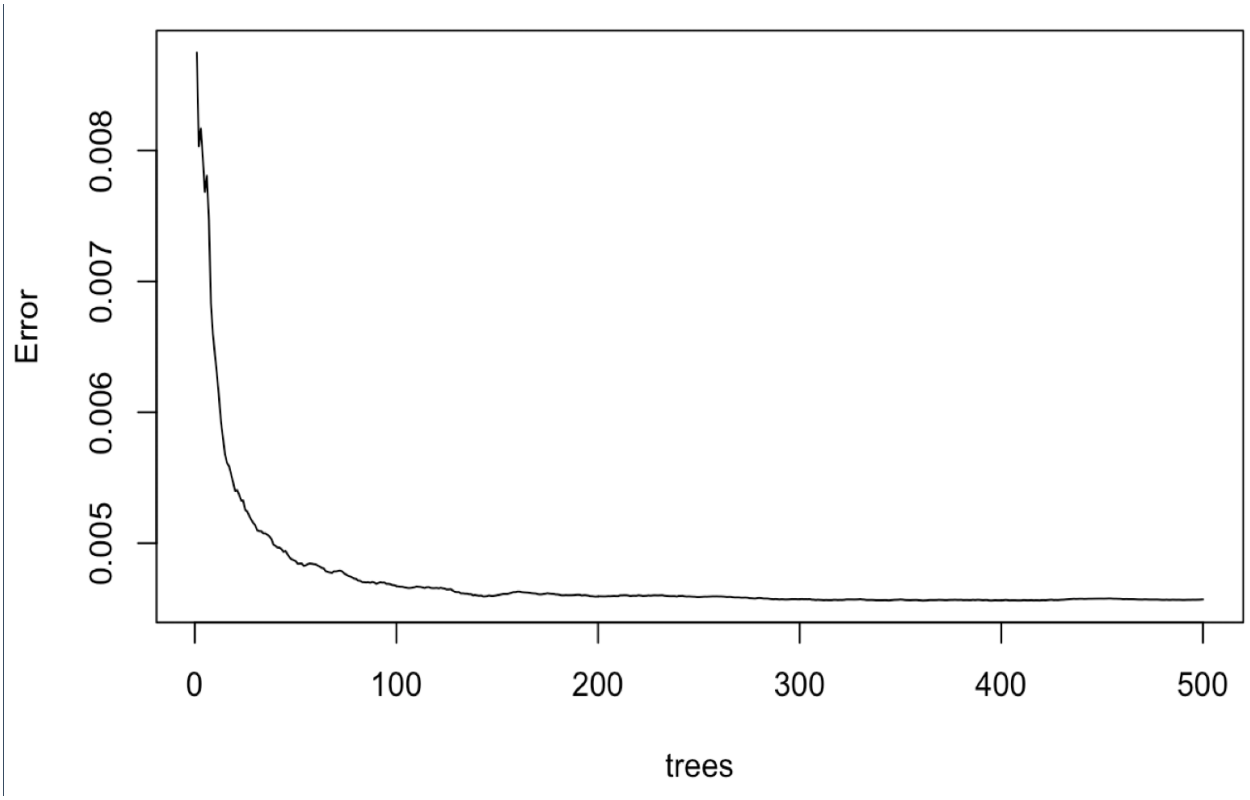
Covariates	**mean(SD), Percentages for dichotomous**
Hemoglobin Difference(first-last)(g/dl)	-1.23(1.29)
Age	57.33(16.25)
Male	53.9%
Ethnicity:	
White	72.1%
Black	11.4%
Hispanic	4.0%
Asian	2.7%
Other	9.1%
Weight (kg)	82.01(24.62)
Creatinine level(mg/dl) first measurement	1.37(1.44)
Urine output (ml/6 hour)	642.85(571.07)
Total Fluid administered(ml)	1446.02(1382.75)
Hemoglobin first measurement	12.64(2.13)
Time difference between first and last hemoglobin measurement	15.64(6.05)
Mechanical ventilation	10.1%
Oxygen Therapy	44.5%

[Table 3: Characteristics of variables after imputation]

Covariates	Mean (SD), Percentages for dichotomous
Creatinine level(mg/dl) first measurement	1.37(1.42)
Urine output (ml/6 hour)	645.00(541.93)
Weight (kg)	81.68(22.27)

Linear regression and stepwise regression model had MSE of 1.08 g/dl [Table 4]. The final model in stepwise regression included variables for total fluid administered, first hemoglobin measurement, time difference between first and last measurement, creatinine level, urine output over 6 hours, weight, gender and all ethnicities except for Hispanic. We determined that 500 trees were enough to establish stable error for random forest model [Figure 4]. Final random forest model had node size of 2, number of variables included at each split were 14 and sample size used was 0.55. MSE for RF model was 1.05 g/dl.

[Figure 4 : Error rate(Y-axis) as we average across number of trees(X-axis)]



[Table 4: Results of Linear regression model]

No.	Term	Estimate	Standard error	Test statistic	P- value
1	(Intercept)	0.48327724	0.02098207	23.0328627	1.89E-102
2	cre_1st	-0.174675	0.02756319	-6.3372556	2.99E-10
3	uo_6hr	0.07751204	0.01802249	4.30085053	1.80E-05
4	mean_wt	0.08127474	0.0222927	3.64580092	2.75E-04
5	age	0.00037988	0.00686112	0.05536673	9.56E-01
6	HEMOGLOBIN_1st	-0.3377439	0.01272451	-26.542778	2.93E-130
7	hr	-0.05983	0.00789137	-7.5817074	5.56E-14
8	MechVent	0.00195736	0.00565402	0.34618891	7.29E-01
9	OxygenTherapy	0.0034501	0.00358851	0.96142826	3.36E-01
10	total_ns_cv	-0.0983803	0.01569095	-6.2698772	4.57E-10
11	Gender	0.01958419	0.00360058	5.43917888	6.13E-08
12	ethnicity_asian	0.02726622	0.02000419	1.36302562	1.73E-01
13	ethnicity_hisp	0.00512841	0.01934268	0.2651344	7.91E-01
14	ethnicity_white	0.0222373	0.01776104	1.25202745	2.11E-01
15	ethnicity_black	0.01939088	0.01838873	1.05449842	2.92E-01
16	ethnicity_other	0.02700241	0.01857319	1.45383817	1.46E-01

The support vector machine model, after tuning, selected a final model where Kernel was radial, C was 7, gamma was 0.005 and epsilon 0.1. The model had 1538 support vectors. MSE for the model was 1.02 g/dl.

The gradient boosting machine model, after tuning, selected a final model that had 1254 trees with interaction depth of 3 nodes, learning rate of 0.01. Minimum number of observations allowed in the trees terminal nodes was 3 and bagging fraction was 0.65. MSE for GBM was1.03 g/dl.

The neural network final model had 5 hidden layers and decay of 0.1. Neural network model had MSE of 1.07 g/dl.

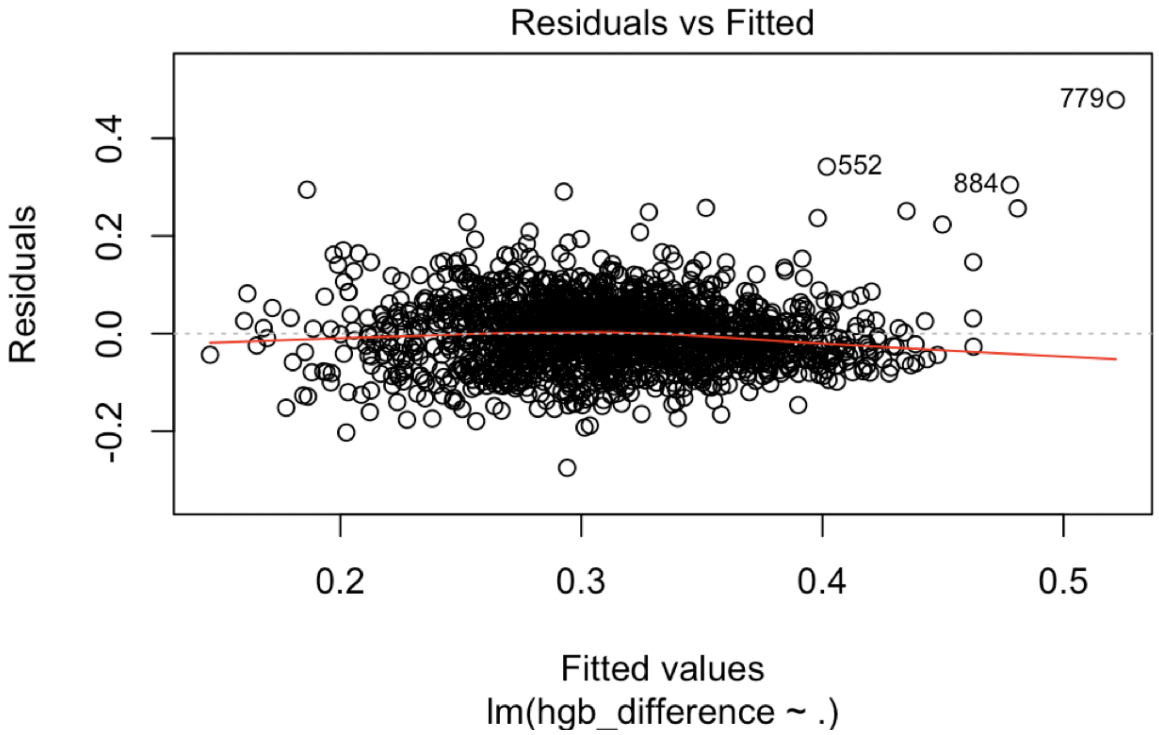
We finalized support vector machine model as our best and final model as it was most parsimonious with the lowest MSE. Other comparative statistics for all 6 models can be found in Table 5.

[Table 5: Model Comparison]

	Average Error	**Mean Squared Error (MSE)**	**Median Error**	**Average Absolute Error**	**Median Absolute Error**
General linear model	-0.01426741	1.081221	-0.01819346	0.8047973	0.6374809
Stepwise regression AIC	-0.01551956	1.083188	-0.005111501	0.8061419	0.6444805
Random Forest Model	-0.0075092	1.045226	-0.0199821	0.7924891	0.6376987
Support Machine Vectors	0.021959	1.015289	0.006805962	0.7784085	0.6270712

Gradient Boosting Machines	-0.00257172	1.027221	-0.02443405	0.7793252	0.6173611
Neural Network	-0.01329463	1.072719	-0.01031218	0.8033177	0.6459622

[Figure 5: Assessing Linear regression model]



Discussion

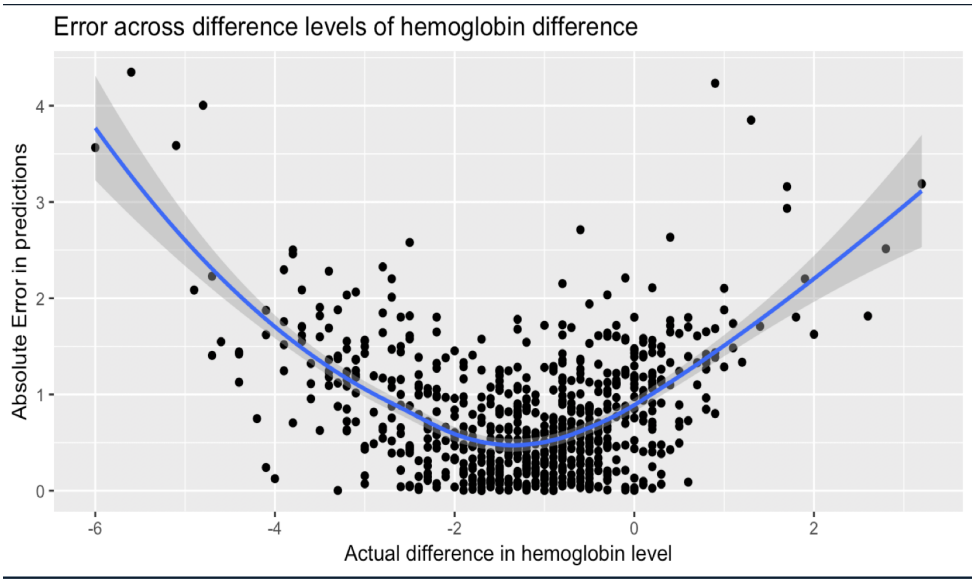
This study had the objective of predicting the hemodilution effect after fluid resuscitation using the ICU data available for critically ill patients. After trying six different models, support vector machine model was finalized as the best model. The final model had an MSE of 1.02g/dl and RMSE of 1.01 g/dl. Results suggest that hemodilution exists in the critically ill cohort who receive normal saline over the first 24 hours. However, even with our best predictive model, the prediction may have an error of 1.01 g/dl on both directions.

In a simple regression model, a discernible pattern can be seen when residuals are plotted against fitted values [Figure 5] and raised concern for linearity of the model. To fit a more non-linear model, we built random forest, support vector machines, gradient boosting machine and neural network models. SVM was finalized as best model as it had best mean square error (1.02), however, for all models MSE ranges from 1.02 to 1.08 g/dl. Support vector machines consider regularization parameters and avoids overfitting of the model. Optimal models with support vector machines are for binary classifiers where the output is a choice between 2 classes.

Comparing SVM to the rest of the models is instructive. Random forest models do not require much data preparation and are very good when using variables both categorical and continuous. They are also robust to the outliers and have better interpretability. Gradient boosting models provide superior predictive accuracy amongst all options. They have lot of flexibility, can handle missing data quite well, and can optimize on different loss functions. They provide several hyperparameter tuning options that make the function fit very flexible. Both SVM and GBM, however, are computationally expensive, and require a lot of processing power. As a whole, they are can be difficult to interpret. Neural networks usually outperform most of the machine learning models. The disadvantage of Neural Networks lies in their "black box" nature. Neural Networks usually require much more data than traditional machine learning algorithms. The sample size of our cohort may actually be too small for an NN model.

In terms of clinical utility of the model, a significant hemorrhage is defined as 2 g/dl decrease in hemoglobin [24], our model predicts with confidence an interval of width of about 2 g/dl. Hence SVM, even though the best model amongst all six, may not be clinically useful. Looking at median errors and MSE, we can say that the model has a lower error rate when the decrease in hemoglobin is around the median decrease in hemoglobin found in our data, and that errors increase in extreme cases [Figure 6].

[Figure 6: Absolute error vs Hemoglobin difference]



In the model, the most influential variable is the first hemoglobin measurement. The interpretation might be that the first measurement represents the overall hemodynamic status of patients while ICU admission. Interestingly, the creatinine level is the second highest influence in the SVM model. It might be explained by the fact that higher creatinine representing poor kidney function which is related to low erythropoietin stimulation, resulting in a decrease in hemoglobin. Loss of kidney function would also result in hypervolemia in the body. We also observed that the duration of ICU stay is a significant factor affecting hemoglobin change. However, this observation might not be directly correlated to the hemodilution effect, as patients might receive intensive and invasive procedures e.g. arterial catheters during the first few hours of ICU admission, a phenomenon referred to as iatrogenic anemia [15]. Ethnicity also has a relatively high significance in our result. In the previous study, a similar hematologic difference of among different ethnicity group [25] was observed.

Our study has several limitations. The cohort design implies that we're not able to consider the dynamic status of disease or any other underlying disease that the patients may have other than coded ICD-9 diagnosis. We are also not able to exclude patients who underwent invasive procedures (such as central lines or arterial catheters) during the measurements. The mismatch of timing between blood draws and normal saline infusion may include more occult confounders. However, if we adjust the tolerated mismatch time to be less than one hour, we would have less than 10% of the patients compared to the current cohort study due to missing data. In terms of laboratory measurement, the Hgb/Hct lab test has a not insignificant margin of error: levels fluctuate depending on site of draw and time of draw, and Hgb/Hct is not homogenous throughout the entire venous plexus. One way to improve our model is to take records of more frequent blood draws, ideally every 4 to 6-hour Hgb/Hct draws over the first 24 hours. However, increased phlebotomy itself may cause more decrease in hemoglobin [25].

In conclusion, our study has proven that hemodilution exists in the critically ill cohort using normal saline over the first 24 hours, but there is either no simple model to predict the expected drop or we are limited by the data. Future studies should further consider refinement of these models and experimentation with new avenues of research to accomplish our initial goal.

Supplement 1

[Model comparison for models using log transformed variables]

	Average Error	Mean Squared Error (MSE)	Median Error	Average Absolute Error	Median Absolute Error
General linear model	0.0201445	1.101263	0.02207029	0.8038478	0.6452676
Stepwise AIC	0.02032311	1.103542	0.02855068	0.804429	0.6385515
Random Forest Model	0.04368445	1.238663	0.03499506	0.8384529	0.6252178
Support Machine Vectors	0.1052783	1.370987	0.09305351	0.8560119	0.6724003
Gradient Boosting Machines	0.0656093	1.350842	0.02839761	0.869207	0.6748819
Neural Network	0.04002857	1.086694	0.0654735	0.8153645	0.6695578

Reference

[1] Schmaier, A. H., & Lazarus, H. M. (2012). Concise guide to hematology. Chichester, West Sussex, UK: Wiley-Blackwell.

[2] Hemodilution. (n.d.). Retrieved from https://www.merriam-webster.com/dictionary/hemodilution

[3] Perel, A. (2017). Iatrogenic hemodilution: A possible cause for avoidable blood transfusions? Critical Care,21(1). doi:10.1186/s13054-017-1872-1

[4]Cervera, A. L., & Moss, G. (1974). Crystalloid Distribution Following Hemorrhage And Hemodilution. The Journal of Trauma: Injury, Infection, and Critical Care,14(6), 506-520. doi:10.1097/00005373-197406000-00007

[5]Wadström, J., & Gerdin, B. (1990). Effect of Bleeding and Hypervolaemic Haemodilution on Traumatic Vasospasm in Rabbits. Scandinavian Journal of Plastic and Reconstructive Surgery and Hand Surgery,24(2), 107-111. doi:10.3109/02844319009004529

[6]Hannon, J., Wade, C., Bossone, C., Hunt, M., Coppes, R., & Loveday, J. (1990). Blood gas and acid-base status of conscious pigs subjected to fixed-volume hemorrhage and resuscitated with hypertonic saline dextran. Resuscitation,20(3), 265-266. doi:10.1016/0300-9572(90)90015-7

[7]Kowalenko, T., Stern, S., Wang, X., Dronen, S., & Hurst, J. M. (1991). Improved Outcome With “Hypotensive” Resuscitation Of Uncontrolled Hemorrhagic Shock In A Swine Model. The Journal of Trauma: Injury, Infection, and Critical Care,31(7), 1032. doi:10.1097/00005373-199107000-00068

[8]Hwang, N. C. (2015). Preventive Strategies for Minimizing Hemodilution in the Cardiac Surgery Patient During Cardiopulmonary Bypass. Journal of Cardiothoracic and Vascular Anesthesia,29(6), 1663-1671. doi:10.1053/j.jvca.2015.08.002

[9] Muraki, R., Hiraoka, A., Nagata, K., Nakajima, K., Oshita, T., Arimichi, M., . . . Sakaguchi, T. (2018). Novel method for estimating the total blood volume: The importance of adjustment using the ideal body weight and age for the accurate prediction of haemodilution during cardiopulmonary bypass. Interactive CardioVascular and Thoracic Surgery. doi:10.1093/icvts/ivy173

[10]Perel, A. (2018). The relationship between the decrease in haemoglobin concentration and the volume of fluids administered during resuscitation from septic shock may not be so “weak”. Critical Care,22(1). doi:10.1186/s13054-018-2118-6

[11]Paydar, S., Bazrafkan, H., Golestani, N., Roozbeh, J., Akrami, A., & Moradi, A. M. (2014). Effects of Intravenous Fluid Therapy on Clinical and Biochemical Parameters of Trauma Patients. Emergency, 2(2), 90–95.

[12]Ross, S. W., Christmas, A. B., Fischer, P. E., Holway, H., Seymour, R., Huntington, C. R., . . . Sing, R. F. (2018). Defining Dogma: Quantifying Crystalloid Hemodilution in a Prospective Randomized Control Trial with Blood Donation as a Model for Hemorrhage. Journal of the American College of Surgeons,227(3), 321-331. doi:10.1016/j.jamcollsurg.2018.05.005

[13] Johnson AEW, Pollard TJ, Shen L, Lehman L, Feng M, Ghassemi M, Moody B, Szolovits P, Celi LA, Mark RG. MIMIC-III, a freely accessible critical care database. Scientific Data (2016).

[14] Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, Ivanov PCh, Mark RG, Mietus JE, Moody GB, Peng C-K, Stanley HE. PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals. Circulation 101(23):e215-e220 [Circulation Electronic Pages; http://circ.ahajournals.org/content/101/23/e215.full]; 2000 (June 13).

[15] Pollard, T. (n.d.). Getting Started. Retrieved from https://eicu-crd.mit.edu/
Lai CW, Starkie T, Creanor S, Struthers RA, Portch D, Erasmus PD, Mellor N, Hosie KB, Sneyd JR, Minto G. Randomized controlled trial of stroke volume optimization during elective major abdominal surgery in patients stratified by aerobic fitness. Br J Anaesth. 2015;115:578–89.

[16] Nguyen BV, Bota DP, Melot C, Vincent JL. Time course of hemoglobin concentrations in nonbleeding intensive care unit patients. Crit Care Med. 2003;31:406–410. doi: 10.1097/00003246-200301001-00001.

[17] Acute Normovolemic Hemodilution Reduces Allogeneic Red Blood Cell Transfusion in Cardiac Surgery: A Systematic Review and Meta-analysis of Randomized Trials.
Barile L1, Fominskiy E, Di Tomasso N, Alpìzar Castro LE, Landoni G, De Luca M, Bignami E, Sala A, Zangrillo A, Monaco F.

[18] Michael Mayer (2018). missRanger: Fast Imputation of Missing Values. R package version 1.0.3. https://CRAN.R-project.org/package=missRanger

[19] Breiman, L. (2001). Random Forests. *Machine Learning,45*(1), 5-32. doi:10.1023/a:1010933404324

[20] Marvin N. Wright, Andreas Ziegler (2017). ranger: A Fast Implementation of Random Forests for High Dimensional Data in C++ and R. Journal of Statistical Software, 77(1), 1-17. doi:10.18637/jss.v077.i01

[21] David Meyer, Evgenia Dimitriadou, Kurt Hornik, Andreas Weingessel and Friedrich Leisch (2018). e1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien. R package version 1.7-0. https://CRAN.R-project.org/package=e1071

[22] Brandon Greenwell, Bradley Boehmke, Jay Cunningham and GBM Developers (2018). gbm: Generalized Boosted Regression Models. R package version 2.1.4. https://CRAN.R-project.org/package=gbm

[23] Max Kuhn. Contributions from Jed Wing, Steve Weston, Andre Williams, Chris Keefer, Allan Engelhardt, Tony Cooper, Zachary Mayer, Brenton Kenkel, the R Core Team, Michael Benesty, Reynald Lescarbeau, Andrew Ziem, Luca Scrucca, Yuan Tang, Can Candan and Tyler Hunt. (2018). caret: Classification and Regression Training. R package version 6.0-81. https://CRAN.R-project.org/package=caret

[24] S. SCHULMAN C. KEARONDefinition of major bleeding in clinical investigations of antihemostatic medicinal products in non-surgical patients. First published: 04 April 2005 https://doi-org.ezp-prod1.hul.harvard.edu/10.1111/j.1538-7836.2005.01204.x

[25] A Population-Based Study of Hemoglobin, Race, and Mortality in Elderly Persons XinQi Dong Carlos Mendes de Leon Andrew Artz YuXiao Tang Raj Shah Denis Evans. The Journals of Gerontology: Series A, Volume 63, Issue 8, 1 August 2008, Pages 873–878, https://doi.org/10.1093/gerona/63.8.873