Received: 21 March 2021 | Revised: 13 September 2021 | Accepted: 16 October 2021

DOI: 10.1111/trf.16739

PATIENT BLOOD MANAGEMENT

TRANSFUSION

From demand forecasting to inventory ordering decisions for red blood cells through integrating machine learning, statistical modeling, and inventory optimization

Na Li[1,2,3] | Donald M. Arnold[2,4,5] | Douglas G. Down[3] | Rebecca Barty[2,6] | John Blake[4,7] | Fei Chiang[3] | Tom Courtney[4] | Marianne Waito[4] | Rick Trifunov[4] | Nancy M. Heddle[2,4]

[1]Department of Community Health Sciences, University of Calgary, Calgary, Alberta, Canada

[2]McMaster Centre for Transfusion Research, Department of Medicine, McMaster University, Hamilton, Ontario, Canada

[3]Department of Computing and Software, McMaster University, Hamilton, Ontario, Canada

[4]Centre for Innovation, Integrated Supply Chain and Analytics, Canadian Blood Services, Ottawa, Ontario, Canada

[5]Department of Medicine, Michael G. DeGroote School of Medicine, McMaster University, Hamilton, Ontario, Canada

[6]Southwest Region, Ontario Regional Blood Coordinating Network, Hamilton, Ontario, Canada

[7]Department of Industrial Engineering, Dalhousie University, Halifax, Nova Scotia, Canada

Correspondence

Na Li, Centre for Health Informatics, TRW Building, 3280 Hospital Drive NW, Calgary, Alberta, T2N 4Z6, Canada.
Email: Na.Li@ucalgary.ca

Funding information

Canadian Blood Services, Grant/Award Number: Blood Efficiency Accelerator Award; Mitacs, Grant/Award Numbers: Accelerate industrial postdoc, IT13639

Abstract

Background: The demand and supply of blood are highly variable over time. Blood inventory management that relies heavily on experience-based decisions may not be adaptive to real demand, leading to high operational costs, wastage, and shortages.

Methods: We combined statistical modeling, machine learning, and optimization methods to develop a data-driven demand forecasting and inventory management strategy for red blood cells (RBCs). We then used the strategy to inform daily blood orders. A secondary semi-weekly (twice per week) ordering strategy was developed to handle the last-mile split delivery problem for blood suppliers, characterized by multi-deliveries to the same location multiple times during a short period of time. Both strategies were evaluated using the TRUST database including all patient data across four hospitals in Hamilton, Ontario.

Results: We identified 227,944 RBC transfusions for 40,787 patients in Hamilton, Ontario from 2012 to 2018. The predicted daily demand from the hybrid demand forecasting model was not significantly different from the actual daily demand (paired *t*-test *p*-value = 0.163); however, the proposed daily ordering quantity from the model was significantly lower than the actual ordering quantity (*p*-value <0.001). The proposed daily ordering strategy reduced inventory levels by 38.4% without risk of shortages, leading to an overall cost reduction of 43.0% (95% confidence interval [CI]: 42.3%, 43.7%) compared with the actual cost. The semi-weekly ordering strategy reduced ordering frequency by 62.6% (95% CI: 61.5%, 63.7%).

Conclusion: The proposed data-driven ordering strategy combining demand forecasting and inventory optimization can achieve significant cost savings for healthcare systems and blood suppliers.

KEYWORDS

blood management, RBC transfusion, transfusion service operations

1 | INTRODUCTION

Blood transfusions, including red blood cells (RBCs), platelets, and plasma, are commonly administered therapeutics in healthcare systems. The demand and supply of blood are highly variable over time, affected by component types and processing methods,[1,2] clinical transfusion guidelines and blood bank practices,[3–7] population characteristics,[8–10] medical procedures and disease categories,[11,12] geographic regions[13–15] and epidemic outbreaks like the COVID-19 pandemic.[16–20]

Canadian Blood Services (CBS) is one of the two blood suppliers in Canada, covering all provinces excluding Québec. The blood supply chain is operated as an integrated network comprised of regional CBS distribution centers supporting numerous surrounding hospitals.[21] CBS provides a daily RBC and platelet inventory status report to hospitals and provincial blood offices.[22] Hospitals request blood products from the regional distribution centers typically for next-day delivery. Hospital blood banks provide retrospective monthly aggregated disposition data to CBS.[23] Furthermore, recipients' demographic and hospitals' inventory information are voluntarily reported and not fully disclosed to CBS or the regional distribution centers.[24] In the current practice, overstocking of blood at hospital blood banks leads to high wastage rates and imprecise planning at CBS[24,25]; for example, wastage rate of platelets is 10%–30%. This problem is compounded by hospital blood banks relying heavily on experience-based decisions for blood inventory management. Although there have been several initiatives aimed at reducing waste, over 10,000 units of RBCs and platelets were wasted in each of the last three years in Ontario, at a cost of over $4.6 million per year.[24]

Each hospital blood bank has its own electronic data collection system. As these systems are not designed for inventory management solutions, there are no analytical functions available for demand trend analysis and forecasting, or inventory planning. Hence, ordering decisions are made from human experience, which cannot adaptively account for variations caused by external factors, resulting in frequent over-ordering and excess inventory to ensure that the demand can be met.[26] Orders from hospital blood banks to CBS are delivered by fax without electronic data records at hospitals.[27] This hinders analytical explorations of historical hospital ordering patterns and order fulfillment rate. Although there are limited data collection functions, blood inventory and disposition data are traceable in electronic medical record (EMR) systems. With appropriate data linkage and pre-processing, data analytical methods can be applied toward improving blood demand and supply management using EMR data. Aiming to reduce over-ordering and excess inventory while ensuring sufficient supply, we describe a novel approach for data-driven RBC demand forecasting and inventory management integrating machine learning, statistical modeling, and inventory optimization using EMR data.

2 | METHODS

This was a quality improvement study using retrospective data from four teaching hospitals in Hamilton, Canada from the years 2012–2018. This study was approved by the Canadian Blood Services Research Ethics Board and Hamilton Integrated Research Ethics Board. The primary objective was to optimize ordering decisions for RBC inventory management (reduce wastage due to expiration, ordering frequency, same-day urgent orders representing shortages, and inventory variability), through integrating RBC demand forecasting and inventory optimization. This study considered the aggregated RBC demand of all hospitals at city level for the following reasons: (a) all hospital blood banks in Hamilton are centrally managed by one Transfusion Medicine team; (b) the aggregated demand of a diverse patient population could be representative of overall Canadian RBC demand.

2.1 | Data collection

Study data were pulled from the TRUST database hosted at the McMaster Centre for Transfusion Research. Variables included RBC inventory and product-related data including: received date, issue date, expiry date, product ABO Rh type, dose, and final disposition. These data were then linked to recipient characteristic data including age, sex, patient ABO Rh type, diagnosis, hospital location, laboratory test result (e.g., hemoglobin, platelet count, mean platelet volume [MPV], red cell distribution width [RDW], immunoglobulins [IgG], international normalized ratio [INR], and creatinine), and surgical procedure, which were used as clinical predictors for demand forecasting model development. The data were preprocessed through a systematic data processing framework[28] to a daily aggregated dataset containing daily demand, inventory level, wastage, and number of patients with different clinical characteristics. Over 200 possible clinical variables were considered in model variable selection. Variables with over 70% missing values (nine variables) were excluded from variable selection, for example abnormal immunoglobulin (IgA, IgM) test results. Methods of imputing missing values for individual variables are based on the clinical definition and the use of the variable.[29–31] For example, if a patient had no daily platelet count result on a given day, we assumed the patient had normal platelet counts on that day. Variables are normalized using the min-max method.[32]

2.2 | A hybrid demand forecasting algorithm

Demand for RBC transfusions is subject to long time dependencies, changing patterns in seasonality, and nonlinear effects of clinical factors. To handle these challenges, a hybrid model combining Seasonal and Trend decomposition using Loess (STL) and eXtreme Gradient Boosting (XGBoost) models was applied. STL is a robust and efficient univariate time series model.[33] The advantages of STL over univariate time series models such as ARIMA, include flexibility to handle different types of seasonality, robust estimation of trend, and seasonal components, the ability to decompose time series with missing values, and fast computation. XGBoost is an efficient machine learning model[34] that features fast computation, flexibility to handle sparsity patterns, and has more regularization parameters to control model complexity.[26] A combination of the two models can integrate their complementary advantages, including handling changing patterns in seasonality and nonlinear dependencies, as well as identifying clinical predictors. The hybrid demand forecasting algorithm started with a time series decomposition of the daily RBC demand using an STL model, after which the STL residuals were forecast with an XGBoost model using a set of clinical predictors. The final forecast demand is the sum of the trend and seasonality components from the STL model and the predictions from the XGBoost model. The hybrid model was used to generate daily RBC demand predictions (Figure 1). Technical details of model development, training, and evaluation can be found in Li et al. 2021.[26]

The hybrid algorithm includes two parameters for the STL model (the trend-cycle window representing the span in lags of the Loess window for trend extraction and the seasonal window representing the span in lags for seasonal extraction) and six parameters for the XGBoost model (learning rate, the maximum depth of a tree, the minimum sum of weights of all observations in a leaf, the fraction of observations to be randomly sampled for each tree, the fraction of columns to be randomly sampled for each tree, and the regularization term on weights). The data from 2012 to 2017 were used for training, and cross-validation was used for hyperparameter tuning. All the hyperparameters were tuned using grid search based on a pre-defined parameter space with the training dataset. The optimal hyperparameters were selected with the minimum Root Mean Square Error (RMSE) using 5-fold cross-validation.[35] The variable selection was processed in an iterative manner using the training dataset based on the variable importance scores calculated from the XGBoost models.[36] Model performance was evaluated with the 2018 data using two accuracy measures: Mean Absolute Percentage Error (MAPE) and RMSE. The model performance of the hybrid model was compared to: a single STL model; an STL + linear regression model; and a long short-term memory (LSTM) model.[37]

2.3 | Data-driven daily ordering strategy for inventory management

To optimize inventory management decisions, a multi-period inventory problem for RBC units was constructed based entirely on data-driven demand estimates from the hybrid demand forecasting model.[26] The assumptions of this inventory optimization problem were: (a) CBS had sufficient supply to fulfill orders (the national order fulfillment rate at CBS was over 98% for RBCs); (b) a shelf life of 42 days for RBC units with a fixed duration of 10 days from the date of blood collection to the date received at hospital blood banks; (c) RBC units were issued using a First-In First-Out (FIFO) withdrawal policy and transfused to ABO Rh identical recipients. Four types of costs were considered: routine delivery (to reduce ordering frequency), inventory holding (to reduce excess inventory), urgent delivery (to reduce shortages), and wastage costs (to reduce wastage due to expiration). The ordering strategy was optimized through a data-driven version of a classical (s, S) policy,[38] which considers inventory level and reorder constraints for controlling the cumulative loss and hedging uncertainty from the demand forecasting model. The inventory target was used as an upper limit to avoid excess inventory due to demand over-estimation, and the reorder level was considered as a lower limit to avoid urgent deliveries due to demand under-estimation. The ordering quantity was generated as follows: when the inventory level was below the reorder level, the order quantity was at least the number of units required to bring the inventory level back to the reorder level, but such that the inventory level did not become greater than the inventory target; if the inventory level was above the reorder level, no order was required. These two constraints, inventory target and reorder level, played an important role in controlling inventory dynamics, see Li et al. 2021[26] for technical details.

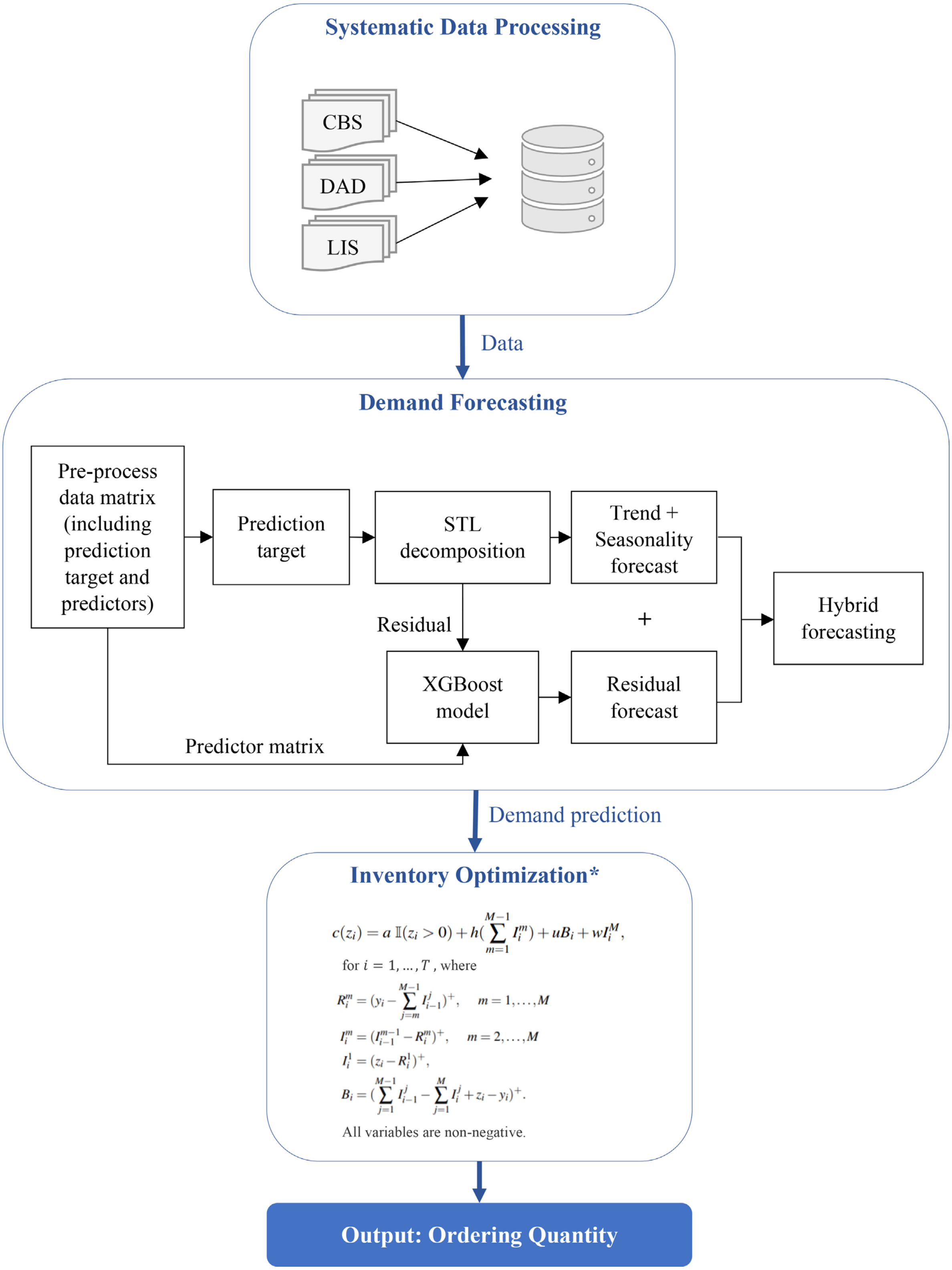


FIGURE 1 The proposed data-driven demand forecasting and inventory management framework (CBS: Canadian Blood Services shipping data; DAD: Discharge Abstract Database; LIS, Laboratory Information System data; STL, Seasonal and Trend decomposition using Loess model; XGBoost, eXtreme Gradient Boosting model. * The optimization problem was defined and solved mathematically in Reference 26) [Color figure can be viewed at wileyonlinelibrary.com]

2.4 | Semi-weekly ordering strategy

Blood suppliers face the last-mile split delivery problem[39]: multiple shipments are delivered to the same location multiple times in a short time frame (e.g., daily). To address this problem, a secondary model with an ordering consolidation strategy was developed through pooling semi-weekly demand. This may improve the transportation economies of scale and reduce the variability for more accurate demand forecasting. The semi-weekly order was defined as the three-day demand for Tuesday through Thursday, and the four-day demand for Friday through Monday. The timing of the orders was fixed to the end of Mondays and Thursdays, and the ordering quantity decision was optimized by the data-driven inventory problem above using the semi-weekly demand estimates from the hybrid demand forecasting algorithm.

2.5 | Descriptive statistical methods for RBC utilization and model evaluation

For descriptive analysis, continuous variables were reported as means and standard deviations (SDs) or medians and interquartile ranges (IQRs). Categorical variables were reported as frequencies and proportions. We presented the descriptive statistics to compare the distributions between the estimated and actual daily demands, as well as between the optimized and actual inventory levels at hospitals. Different measures were compared in pairs between the results based on the actual inventory data and the results after applying the integrated data-driven policy, including the percentage of days with regular deliveries, the percentage of days requiring urgent same-day deliveries, the number of days on hand (DOH), the number of units wasted, and the cost reduction relative to current practice. To emphasize the importance of the inventory constraints (inventory target and reorder level), we compared the inventory levels between ordering entirely based on daily demand predictions and the proposed daily ordering strategy with optimized inventory constraints. To evaluate the sensitivity of different demand forecasting models, we compared these inventory management measures for ordering strategies generated from four demand forecasting models including our proposed STL + XGBoost model, an STL model, an STL + linear regression model, and an LSTM model. We also demonstrated the flexibility of the proposed algorithm by providing an example with a customized mean inventory level (see the Data S1).

3 | RESULTS

The study identified 227,944 RBC transfusions for 40,787 patients in Hamilton, Ontario from 2012 to 2018. The median patient age at transfusion was 67 (IQR: 53, 77) years. 101,837 (44.7%) units were transfused to female recipients (Table 1). 60,982 (26.8%) units were transfused

TABLE 1 RBC utilization and patient demographics

	Statistics
Total number of RBC transfusions	N = 227,944
Age at transfusion (years; median, IQR)	67 (53, 77)
Female (n, %)	101,837 (44.7)
RBC transfusions by recipient ABO group (n, %)	
Group A	87,875 (38.6)
Group O	101,978 (44.7)
Group B	28,152 (12.4)
Group AB	9939 (4.4)
Number of daily RBC transfusions (median, IQR)	93 (70, 112)
Number of ABO non-identical transfusions daily (median, IQR)	5 (3, 8)
Number of Rh-negative units to Rh-positive recipients daily (median, IQR)	7 (4, 10)
Number of units transfused daily to different patient types (median, IQR)	
Patients in medical ICU	23 (18, 30)
Cardiovascular surgery patients	9 (5, 14)
Trauma patients	3 (1, 10)
Hematology/oncology patients	6 (4, 9)
Outpatients	27 (14, 38)
Number of daily transfusions per patient (median, IQR)	3 (2, 5)
Per inpatient	2 (1, 4)
Per outpatient	0 (0, 1)
Laboratory test results prior to transfusions (mean, SD)	
Hemoglobin (g/L)	78.88 (15.90)
Platelet count ($\times 10^9$/L)	173.23 (137.37)
Mean platelet volume (MPV; f/L)	9.24 (1.56)
Red cell distribution width (RDW; %)	16.94 (3.33)
Immunoglobulins (IgG; g/L)	12.15 (12.20)
International normalized ratio (INR)	1.35 (1.05)
Creatinine (μmol/L)	130.94 (130.21)

Abbreviation: IQR: interquartile range; SD: standard deviation.

to patients in medical ICU, 62,316 (27.3%) units to outpatients, 25,436 (11.2%) units to cardiovascular surgery patients, 5460 (2.4%) units to trauma patients, 16,125 (7.1%) units to hematology/oncology patients, 10,510 (4.6%) to emergency department patients, 10,168 (4.5%) to other surgery patients, and 36,947 (16.2%) to other patients. 51,411 (22.6%) of the RBC transfusions were Rh-negative units, and 18,410 (35.8%) of the Rh-negative units were transfused to Rh-positive recipients.

At the participating hospital blood banks, the mean (SD) of daily RBC demand was 92.43 (28.27) units. The mean (SD) of the daily ordering quantity was 103.71 (69.49) units, which was significantly higher than the actual daily demand (difference in means: 11.27; 95% CI: 9.09, 13.46; p-value <0.001). The distribution of the actual daily ordering quantity contained greater variability than the distribution of the actual daily demand (Figure 2). The mean (SD) of monthly expired units was 21.14 (7.69) units, representing an average cost of $8921.08 per month (or $107,053 per year; expiry cost was $422 per unit based on personal communication with CBS). During the period from April 1st, 2012 to May 31st, 2015 (1012 days), CBS reported that 391 (38.6%) days had same-day urgent orders (Table 2).

3.1 | Predicted daily demand versus actual daily demand

After the iterative variable selection process, the final predictors with variable importance from the hybrid model are shown in Figure 3. The mean (SD) of the predicted daily demand from the hybrid model was 92.77 (23.94) units, compared to the actual daily demand (Table 2). The difference in means between the actual and predicted daily demands, −0.34 (95% CI: −0.86, 0.19; p-value = 0.163), was not significant. Table 3 presents the performance of the proposed hybrid demand forecasting algorithm, a single STL model, an STL + linear regression model, and an LSTM model. The accuracy of the proposed hybrid model (MAPE: 15.9%) was significantly higher than the single STL model (MAPE: 25.6%; the lower the MAPE value, the more accurate the model). This reflects the importance of including clinical indicators. The accuracy decreased after replacing XGBoost with a linear regression (MAPE: 17.5%), indicating that XGBoost can better handle the nonlinear patterns in the data. There was no significant improvement of the LSTM model (MAPE: 16.8%). However, the proposed hybrid model requires less computational resources and execution time (hours to train an LSTM model as opposed to minutes for an XGBoost model).

3.2 | Proposed data-driven ordering strategy versus actual practice

The proposed daily ordering quantity was significantly lower than the actual historical ordering quantity (Table 2), reduced by on average 11.27 units (95% CI: 8.89, 13.65; p-value <0.001). Furthermore, there was no significant difference between the actual daily demand and the optimized ordering quantity (difference in means: −0.008; 95% CI: −0.91, 0.89; p-value = 0.987). The inventory level resulting from the proposed policy was significantly lower than the actual inventory level (Figure 4). The mean and SD of DOH was reduced from 13.74 (1.39) to 8.47 (0.46) days. There were no

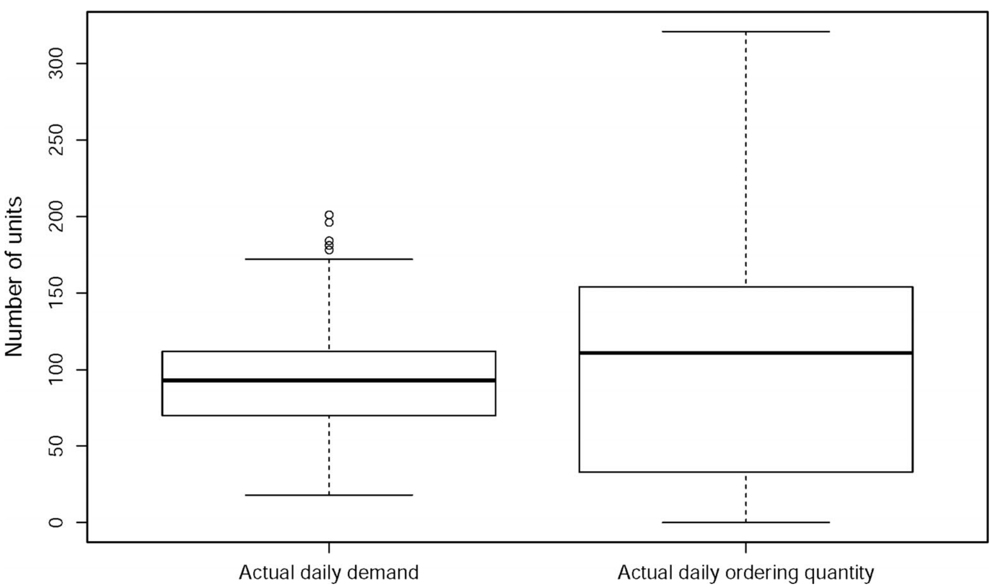


FIGURE 2 Boxplot of daily demand and ordering quantity at hospital blood banks

TABLE 2 Comparisons of different measures between the actual practice and the optimized policy

	Actual practice	Optimized policy	Difference (95% CI)
Initial inventory level, I_0[a]	–	780	–
Daily ordering strategy			
Daily RBC demand (mean, SD)	92.43 (28.27)	92.77 (23.94)	−0.34 (−0.86, 0.19)
Daily ordering quantity (mean, SD)	103.71 (69.49)	92.44 (27.78)	11.27 (8.89, 13.65)
Daily inventory level (mean, SD)	1270.29 (142.66)	782.03 (27.29)	488.26 (482.53, 493.98)
Days of inventory on hand (mean, SD)	13.74 (1.39)	8.47 (0.46)	5.26 (5.20, 5.32)
Percent of days with orders per month (mean, SD)	99.11 (1.78)	96.93 (2.59)	2.18 (1.48, 2.89)
Percent of days with same-day urgent order[b]	38.64	0	38.64 (35.64, 41.72)
Number of units wasted per month (mean, SD)	21.14 (7.69)	0.21 (1.89)	20.93 (19.34, 22.51)
Percent of days with wastage per month (mean, SD)	38.58 (9.51)	0.04 (0.36)	38.54 (36.45, 40.62)
Relative cost reduction over the actual cost (%)			42.96 (42.25, 43.68)
Semi-weekly ordering strategy			
Daily RBC demand based on semi-weekly ordering strategy (mean, SD)[c]	92.36 (147.86)	92.58 (147.64)	−0.22 (−0.59, 0.15)
Daily ordering quantity based on semi-weekly ordering strategy (mean, SD)[c]	103.66 (169.17)	92.39 (141.87)	11.27 (9.38, 13.16)
Daily inventory level (mean, SD)	1270.29 (142.66)	802.10 (86.83)	468.19 (461.57, 474.81)
Days of inventory on hand (mean, SD)	13.74 (1.39)	8.70 (1.02)	5.04 (4.97, 5.11)
Percent of days with orders per month (mean, SD)	99.11 (1.78)	36.51 (5.10)	62.61 (61.47, 63.74)
Number of units wasted per month (mean, SD)	21.14 (7.69)	0.80 (5.94)	20.33 (18.62, 22.05)
Percent of days with wastage per month (mean, SD)	38.58 (9.51)	0.08 (0.50)	38.50 (36.41, 40.59)
Relative cost reduction over the actual cost (%)			45.16 (44.39, 45.92)

[a]The initial inventory level, I_0, is a parameter used to generate the ordering strategies. The mathematical definition was described in Reference 26. It was assumed to be 780 units based on the mean inventory level of the first 3 months in 2008 using Hamilton hospital blood bank inventory data.

[b]Data only available during the period from April 1st, 2012 to May 31st, 2015 (1012 days), so the percentage was calculated by 391/1012. No same-day urgent orders resulted from both the optimized daily and semi-weekly strategies. This comparison was omitted for the semi-weekly strategy.

[c]The semi-weekly demand and ordering quantity were averaged to daily mean and SD. Due to many days including zero semi-weekly demand and ordering quantity, the SDs were much larger than the SDs of the demand and ordering quantity calculated based on daily numbers.

same-day urgent orders and only 17 units expired prior to usage; whereas for the current practice, 38.6% of days required same-day urgent delivery, and 1734 units were expired during the entire period. This can save 20.93 units ($422 per unit) on average per month from outdating and costs for urgent deliveries. The proposed daily ordering policy resulted in a relative cost reduction of 43.0% (95% CI: 42.3%, 43.7%) over current practice.

The proportion of days with orders was reduced by 62.6% by using the proposed semi-weekly ordering strategy. All other measures were similar to the daily ordering strategy (Table 2). The proposed semi-weekly ordering policy resulted in a relative cost reduction of 45.2% (95% CI: 44.4%, 45.9%) over current practice.

3.3 | Sensitivity analysis

Figure 5 compares the inventory levels produced by the proposed algorithm with and without inventory constraints, demonstrating that the inventory constraints can effectively reduce the mean and SD of the inventory level closer to the gold standard (daily ordering by actual demand).

The results for the inventory target and reorder level calculated according to the optimization procedures in Li et al. 2021[26] for four demand forecasting models are shown in Table 3. The inventory targets were very different among the models whereas the reorder levels were similar. The ordering strategy from the proposed hybrid model achieved the highest relative cost reduction and

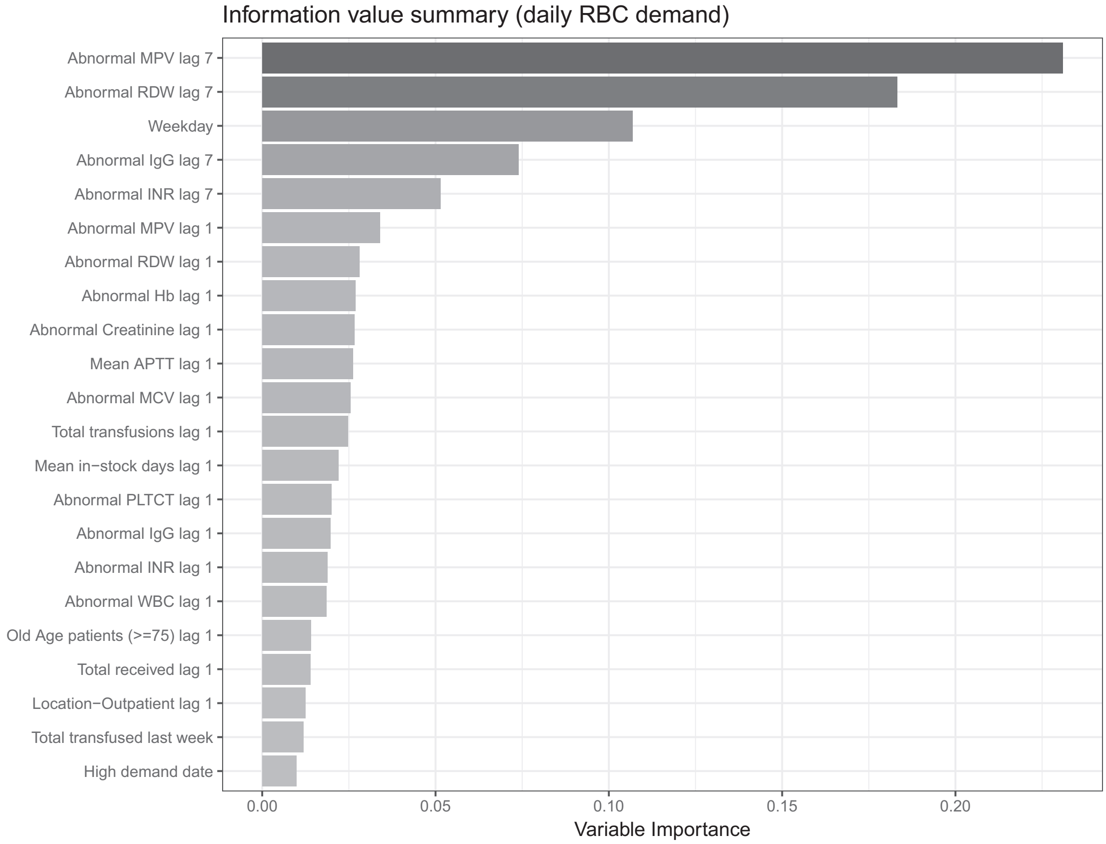


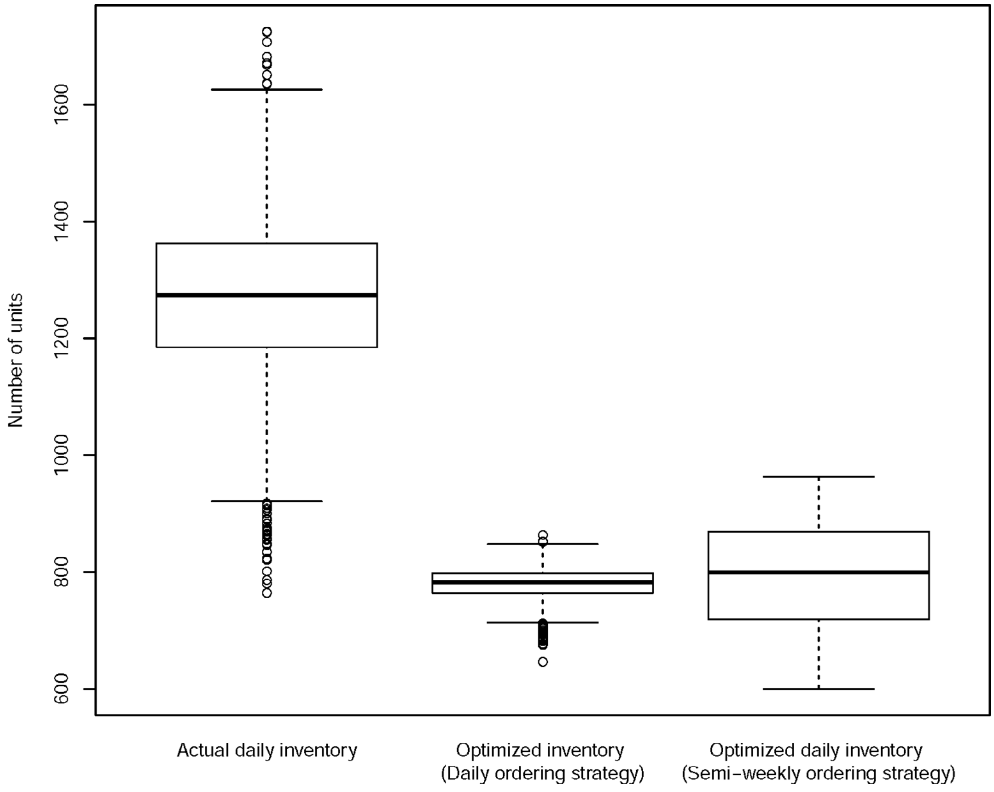
FIGURE 3 Variable importance from the XGBoost model for daily RBC demand. (a) Abnormal represents number of patients with abnormal laboratory test results. (b) Lag 1 represents previous day, and lag 7 represents seven days ago. For example, "Abnormal MPV lag 7" represents number of patients with abnormal MPV results seven days ago. All variables except day of week and high demand date were lagged values since the prediction target was today's demand under the assumption that one would not know any information prior to making a blood order. (c) High demand date is defined as the most frequent dates in a year with demand outliers (i.e., demand lying outside of 1.5 times of the interquartile range). The definitions of the other variables are shown in Table A.1 in Reference 26

TABLE 3 Demand forecasting model performance and inventory measure comparison

	STL + XGBoost	STL	STL + linear regression	LSTM
Model evaluation on test data (2018)				
RMSE	18.32	26.50	19.98	21.44
MAPE (%)	15.9	25.6	17.5	16.8
Calculated inventory constraints from the data-driven inventory optimization problem				
Inventory target	1030	1010	1430	920
Reorder level	830	840	840	840
Daily ordering strategy comparison				
Daily demand prediction (mean, SD)	92.77 (23.94)	92.81 (13.58)	92.23 (23.79)	92.31 (22.86)
Daily ordering quantity (mean, SD)	92.44 (27.78)	92.46 (23.69)	92.43 (26.98)	92.42 (23.56)
Daily inventory level (mean, SD)	782.03 (27.29)	786.39 (35.09)	787.00 (28.57)	785.53 (27.90)
Percent of days with orders per month (mean, SD)	96.93 (2.59)	96.42 (2.88)	97.07 (2.76)	98.51 (2.25)
Relative cost reduction over the actual cost (%)	42.96	42.64	42.56	42.56

Abbreviations: LSTM, long short-term memory model; MAPE, mean absolute percentage error; RMSE, root mean square error; STL, Seasonal and Trend decomposition using Loess model; XGBoost, eXtreme Gradient Boosting model.
Note: The percentage of days with same-day urgent order and the number of units wasted per month were the same for the ordering strategies using the demand forecasts from the models above.

FIGURE 4 Boxplot of daily inventory level



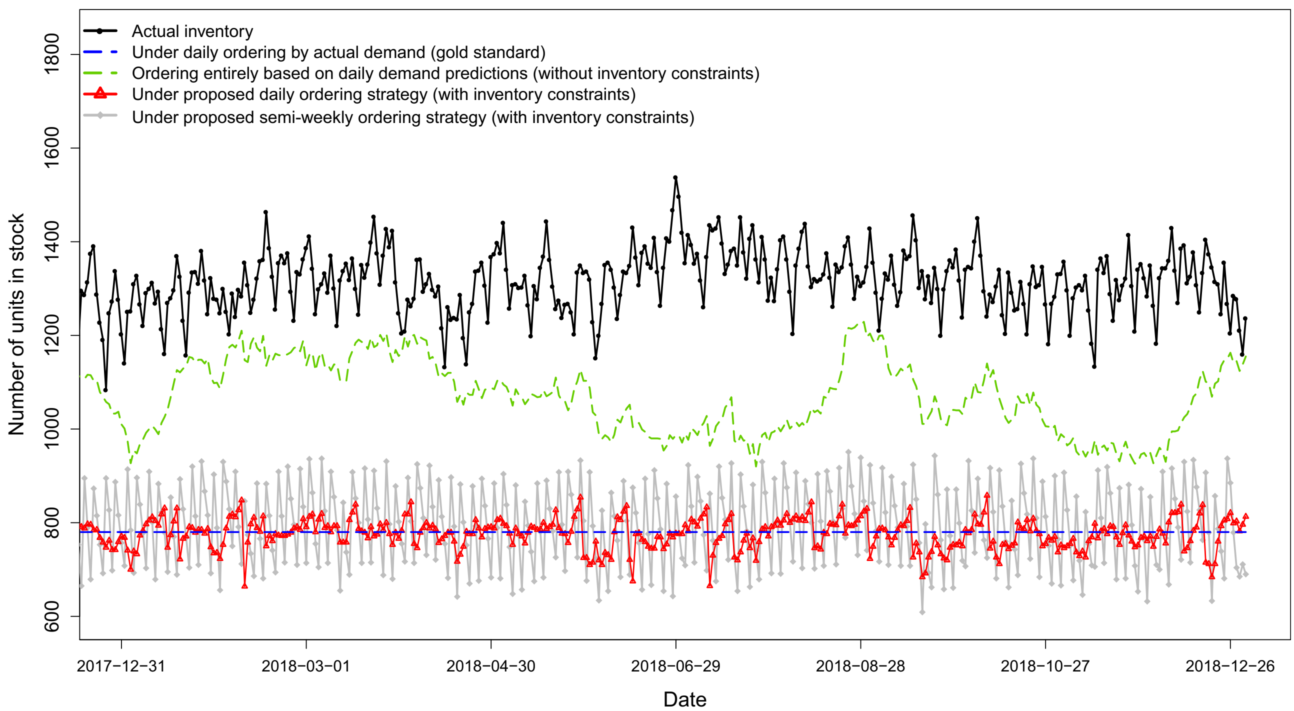


FIGURE 5 Inventory level comparisons among the actual inventory from historical blood inventory data (black line), the constant inventory level assuming ordering by actual demand (gold standard – dashed blue line), the inventory level when ordering by the daily demand predictions from the hybrid algorithm (dashed green line), the inventory level by the proposed daily ordering strategy using both the demand predictions and inventory constraints (red line), and the inventory level by the proposed semi-weekly ordering strategy using both the demand predictions and inventory constraints (gray line) [Color figure can be viewed at wileyonlinelibrary.com]

the lowest mean daily inventory level with the smallest variation, although after applying the inventory constraints the differences were fairly small.

4 | DISCUSSION

In this study, we introduced a novel data-driven blood demand forecasting and inventory management methodology. The hybrid demand forecasting model was highly efficient and achieved the same level of prediction performance as a more complex LSTM model. The hybrid model's daily demand prediction was fed into a multi-period inventory optimization problem to determine daily ordering quantities. We also developed a semi-weekly ordering strategy to consolidate orders into fewer deliveries. Using data-driven demand forecasting and inventory management, the proposed ordering quantity was close to the actual demand. A significant cost saving was achieved, mainly through leaner inventory, contrasting with actual hospital blood bank inventory levels that are excessive, highly variable, and associated with an increasing trend over time.[26]

Although there have been many studies in the field of blood demand and supply management,[40–44] these typically assume independent and identically distributed demands, which is rarely the case, as the demand of blood has temporal dependence and is affected by clinical indicators. To our knowledge, our study team was the first to consider integrating demand forecasting models into blood inventory management strategies using electronic medical records.[26] This enables data-driven decision-making by simultaneously considering clinical indicators and historical demand patterns. It provides an adaptive solution that can help resolve practical challenges for RBC demand and supply chain management. Furthermore, the proposed algorithm is generalizable because of its adaptiveness to data. The demand predictions and inventory constraints can be calculated based on historical inventory data that are not limited to the data from Hamilton hospitals. As shown in Table 3, the proposed algorithm can also produce reasonable ordering decisions based on demand predictions from a univariate time series model without considering clinical predictors, if one is willing to tolerate moderately higher inventory variation than the ordering strategies generated based on multivariate demand forecasting models. The proposed algorithm can also be generalized for other blood products such as platelets.

Through the demand forecasting model, we identified 22 clinical and operational predictors for daily RBC utilization considering various time lags to account for autocorrelations (Figure 3). Among these predictors, day of week was identified as the most important operational predictor. Weekday/weekend effects are commonly observed in blood demand data.[25,26,45] Numbers of patients with abnormal MPV, RDW, IgG, and INR at lag 7 were the top four clinical predictors. These top predictors are consistent with other work in the area, see Reference 45. As this study was aimed at RBC inventory management, the biological plausibility of key predictors was not directly addressed. Recent publications on the prediction of perioperative RBC transfusions[46,47] and the prediction of RBC transfusion for patients with acute gastrointestinal bleeding[48] using patient level data also showed that laboratory test results (e.g., INR, creatinine, RDW, etc.) were important predictors for RBC transfusions. We hypothesize the predictors captured by the XGBoost model indirectly predict the need of preoperative RBC transfusions and transfusions for patients who developed or were in presence of inflammatory conditions, anemia, or with active/acute bleeding.[1,49–52] A more systematic approach to evaluate biological plausibility is required to test our hypothesis, which is beyond the scope of this study. The 7-day time lag reflects the temporal relationship associated between the disease severity or symptom development and the clinical decision process for an RBC transfusion. Although abnormal hemoglobin is known to be the primary trigger of RBC transfusions, the duration from an abnormal hemoglobin result to a transfusion decision is usually in hours, occurring on the same day. Since the model was aimed to predict next-day RBC demand, same-day hemoglobin results cannot be used as a predictor. Hence, abnormal hemoglobin at lag 1 (one day prior to the demand target) was identified as a predictor, but with less importance than the abnormal laboratory results stated above.

There are three limitations of the proposed methodology. First, the inventory optimization problem assumes RBC units were issued to ABO Rh identical recipients at hospital blood banks. This means that for each blood group, the daily (or semi-weekly) order was calculated by the percentage of blood group multiplying the ordering decision for all blood groups from the model. This assumption may cause issues in practice. The second limitation is the assumption of a fixed storage duration of 10 days at CBS. We found a significant increasing trend of the storage duration after 2017. This could be associated with a potential increase of the number of blood donations due to the online appointment booking system that was launched at CBS in 2017. If the increase in blood supply exceeded the demand, it can result in a longer storage duration at CBS. However, the increasing trend of storage duration at CBS also addresses the need for a better blood inventory management strategy to support blood donation collection planning for blood suppliers.

Third, the proposed semi-weekly strategy may not be applicable to all hospital blood banks, especially for hospital blood banks with very limited inventory capacity or hospitals with sparse demand, since this strategy may result in higher inventory variation leading to wastages or shortages. Furthermore, this strategy would require certain operational procedure revisions and/or additional training. For example, hospital staff would need to be prepared for a much larger delivery than the current daily order.

Data accessibility could present a challenge. For small or rural hospitals, certain laboratory tests, such as MPV and IgG, may not be available. Our suggestion is to train site-specific models with available EMR, so that the model can select the predictors that are most reflective of the site's demands and produce a corresponding ordering strategy. In the situation where no EMR data are available, we have shown that using a simple univariate time series model based on the historical number of RBC transfusions alone can also achieve significant improvement (Table 3). In order to apply the methodology, blood bank staff should consult a data scientist to advise the best use of data and methodology. To simplify the process and increase knowledge mobilization, we will design an easy-to-use software tool for blood bank implementation. With such a tool, blood bank staff would be able to generate predicted demands and ordering quantities in seconds, with easy integration into their routine ordering practice.

As next steps, we plan to (a) further investigate optimal ordering strategies by ABO Rh types; (b) apply the proposed algorithm to hospital blood bank data in other cities and provinces, as well as other blood products, (c) design an interactive platform to automatically generate ordering strategies using the proposed methodology, and (d) implement the proposed methodology at hospital blood banks.

ACKNOWLEDGMENTS

This study was funded by Mitacs through the Accelerate Industrial Postdoc program (Grant Number: IT13639) in collaboration with Canadian Blood Services. The funding support from Canadian Blood Services was through the Blood Efficiency Accelerator program, funded by the federal government (Health Canada) and the provincial and territorial ministries of health. The views herein do not necessarily reflect the views of Canadian Blood Services or the federal, provincial, or territorial governments of Canada.

ORCID

Na Li https://orcid.org/0000-0002-4803-0984
Donald M. Arnold https://orcid.org/0000-0003-0943-8853
Rebecca Barty https://orcid.org/0000-0002-9024-6156
John Blake https://orcid.org/0000-0003-0617-8996
Nancy M. Heddle https://orcid.org/0000-0002-9671-3501

REFERENCES

1. Sharma S, Sharma P, Tyler LN. Transfusion of blood and blood products: indications and complications. Am Fam Physician. 2011;83(6):719–24.
2. Singh VK, Saini A, Tsuji K, Sharma PB, Chandra R. Manufacturing blood ex vivo: a futuristic approach to deal with the supply and safety concerns. Front cell. Dev Biol. 2014;2 (June):1–26.
3. Carson JL, Guyatt G, Heddle NM, Grossman BJ, Cohn CS, Fung MK, et al. Clinical practice guidelines from the AABB: red blood cell transfusion thresholds and storage. J Am Med Assoc. 2016;316(19):2025–35.
4. Storch EK, Custer BS, Jacobs MR, Menitove JE, Mintz PD. Review of current transfusion therapy and blood banking practices. Blood Rev. 2019;38:100593.
5. Kaufman RM, Djulbegovic B, Gernsheimer T, Kleinman S, Tinmouth AT, Capocelli KE, et al. Platelet transfusion: a clinical practice guideline from the AABB. Ann Intern Med. 2015; 162(3):205–13.
6. Schiffer CA, Bohlke K, Delaney M, Hume H, Magdalinski AJ, McCullough JJ, et al. Platelet transfusion for patients with cancer: American society of clinical oncology clinical practice guideline update. J Clin Oncol. 2018; 36(3):283–99.
7. National Institute for Clinical Care Excellence Blood transfusion | guidance and guidelines|NICE. Natl Inst Heal Care Excell. 2015;(November): Available from: https://www.nice.org.uk/guidance/ng24/chapter/Recommendations
8. Vamvakas EC. Epidemiology of blood transfusion. Transfusion. 1994;34(6):464–70.
9. Kleinman S, Busch MP, Murphy EL, Shan H, Ness P, Glynn SA. The national heart, lung, and blood institute recipient epidemiology and donor evaluation study (REDS-III): a research program striving to improve blood donor and transfusion recipient outcomes. Transfusion. 2014;54(3 Pt 2): 942–55.
10. Williamson LM, Devine DV. Challenges in the management of the blood supply. Lancet. 2013;381:1866–75.
11. Zewdie K, Genetu A, Mekonnen Y, Worku T, Sahlu A, Gulilalt D. Efficiency of blood utilization in elective surgical patients. BMC Health Serv Res. 2019;19(1):804.
12. Roberts N, James SL, Delaney M, Fitzmaurice C. Blood transfusion trends by disease category in the United States, 2000 to 2014. Transfus Apher Sci. 2021;60(1):103012.
13. Mammen JJ, Asirvatham ES. The demand and supply of blood in India. Lancet Haematol. 2020;7(2):e94.
14. Roberts N, James S, Delaney M, Fitzmaurice C. The global need and availability of blood products: a modelling study. Lancet Haematol. 2019;6(12):e606–15.
15. Bosch MA, Contreras E, Madoz P, Ortiz P, Pereira A, Pujol MM. The epidemiology of blood component transfusion in Catalonia, northeastern Spain. Transfusion. 2011;51(1): 105–16.

16. Stanworth SJ, New HV, Apelseth TO, Brunskill S, Cardigan R, Doree C, et al. Effects of the COVID-19 pandemic on supply and use of blood for transfusion. Lancet Haematol. 2020;7(10): e756–e764.
17. Govindan K, Mina H, Alavi B. A decision support system for demand management in healthcare supply chains considering the epidemic outbreaks: a case study of coronavirus disease 2019 (COVID-19). Transp Res Part E Logist Transp Rev. 2020; 138(101967):1–14.
18. Satyavarapu A & Wagle D Improving the fragile US supply of blood. McKinsey & Company. 2020;7 Available from: https://www.mckinsey.com/industries/public-and-social-sector/our-insights/improving-the-fragile-us-supply-of-blood
19. Canadian Blood Services. Rising demand for blood ushers in National Blood Donor Week, June 8–14 2020. Available from: https://www.blood.ca/en/about-us/media/newsroom/rising-demand-for-blood-ushers-in-national-blood-donor-week-june-8-14
20. The Canadian Press. Increased demand from surgeries causing concern for Canadian Blood Services. 2020. Available from: https://www.ctvnews.ca/health/increased-demand-from-surgeries-causing-concern-for-canadian-blood-services-1.4974862
21. Canadian Blood Services. Inventory planning and management. Inventory Best Practices. 2015;147–83. Available from: https://www.blood.ca/en/hospital-services/inventory-ordering/inventory-planning-and-management
22. Canadian Blood Services. Daily red cell and platelet inventory status report. 2010. Available from: https://www.blood.ca/sites/default/files/Report-Legend.pdf
23. Canadian Blood Services. Blood component and product disposition system. 2019. Available from: https://www.blood.ca/sites/default/files/Disposition-Inventory-UserGuide-Version_3.0.1.2.pdf
24. Office of the Auditor General of Ontario. Blood management and safety. Value-for-money audit. 2020;1–68. Available from: https://www.auditor.on.ca/en/content/annualreports/arreports/en20/20VFM_02bloodmgmt.pdf
25. Motamedi M, Li N, Down DG & Heddle NM. Demand forecasting for platelet usage: from univariate time series to multivariate models. arXiv.org. Preprint. Available from: http://arxiv.org/abs/2101.02305
26. Li N, Chiang F, Down DG, Heddle NM A decision integration strategy for short-term demand forecasting and ordering for red blood cell components. Operations Research for Health Care. 2021;29100290. https://doi.org/10.1016/j.orhc.2021.100290
27. Canadian Blood Services. Submitting product orders. Available from: https://www.blood.ca/en/hospital-services/inventory-ordering/submitting-product-orders
28. Tang S, Davarmanesh P, Song Y, Koutra D, Sjoding MW, Wiens J. Democratizing EHR analyses with FIDDLE: a flexible data-driven preprocessing pipeline for structured clinical data. J Am Med Inform Assoc. 2020;27(12):1921–134.
29. Purwar A, Singh SK. Hybrid prediction model with missing value imputation for medical data. Exp Syst Appl. 2015;42(13): 5621–31.
30. Rahman MM, Davis DN. Machine learning-based missing value imputation method for clinical datasets. IAENG Transactions on Engineering Technologies; Switzerland: Springer Nature Switzerland AG; 2013. p. 245–57.
31. Khan SI, Hoque ASML. SICE: an improved missing data imputation technique. J Big Data. 2020;7(1):37.
32. Manimekalai K, Kavitha A. Missing value imputation and normalization techniques in myocardial infarction. ICTACT Journal on Soft Comput. 2018;8(3):1655–62.
33. Cleveland RB, Cleveland WS, McRae JE, Terpenning I. STL: a seasonal-trend decomposition procedure based on loess (with discussion). J Off Stat. 1990;6:3–73. Available from: http://cs.wellesley.edu/~cs315/Papers/stlstatisticalmodel.pdf
34. Chen T, Guestrin C. XGBoost: a scalable tree boosting system. Proceedings of the ACM SIGKDD international conference on knowledge discovery and data mining; 2016. p. 785–94.
35. Kohavi R. A study of cross-validation and bootstrap for accuracy estimation and model selection. Proceedings of the 14th International Joint Conference on Artificial Intelligence - Volume 2; 1995. p. 1137–43.
36. Ruppert D. The elements of statistical learning: data mining, inference, and prediction. J Am Stat Assoc. 2004;99(466):567.
37. Zia T, Zahid U. Long short-term memory recurrent neural network architectures for Urdu acoustic modeling. International Journal of Speech Technology. 2019;22:21–30.
38. Scarf H. The optimality of (S,s) policies in the dynamic inventory problem. Mathematical Methods in Social Science; 1960. p. 196–202.
39. Zhang Y, Sun L, Hu X, Zhao C. Order consolidation for the last-mile split delivery in online retailing. Transportation Research Part E: Logistics and Transportation Review. 2019; 122:309–27.
40. Mansur A, Vanany I, Indah Arvitrida N. Challenge and opportunity research in blood supply chain management: a literature review. MATEC Web of Conferences 2018;154:01092. https://doi.org/10.1051/matecconf/201815401092
41. Sirelson V, Brodheim E. A computer planning model for blood platelet production and distribution. Comput Methods Programs Biomed. 1991;35(4):279–91.
42. Sarhangian V, Abouee-Mehrizi H, Baron O, Berman O. Threshold-based allocation policies for inventory management of red blood cells. Manuf Serv Oper Manag. 2018;20(2):347–62.
43. Attari MYN, Pasandideh SHR, Aghaie A, Niaki STA. A bi-objective robust optimization model for a blood collection and testing problem: an accelerated stochastic Benders decomposition. Annals of Operations Research 2018;1–39. https://doi.org/10.1007/s10479-018-3059-9
44. Yousefi Nejad Attari M, Pasandideh SHR, Akhavan Niaki ST. A hybrid robust stochastic programming for a bi-objective blood collection facilities problem (case study: Iranian blood transfusion network). J Ind Prod Eng. 2019;36(3):154–67.
45. Guan L, Tian X, Gombar S, Zemek AJ, Krishnan G, Scott R, et al. Big data modeling to predict platelet usage and minimize wastage in a tertiary care system. Proc Natl Acad Sci U S A. 2017;114(43):11368–11373.
46. Walczak S, Velanovich V. Prediction of perioperative transfusions using an artificial neural network. PLoS One. 2020;15(2): e0229450.
47. Huang X, Wang Y, Chen B, Huang Y, Wang X, Chen L, et al. Ability of a machine learning algorithm to predict the need for perioperative red blood cells transfusion in pelvic fracture patients: a multicenter cohort study in China. Front Med. 2021;8: 1–12. https://doi.org/10.3389/fmed.2021.694733
48. Shung D, Huang J, Castro E, Tay JK, Simonov M, Laine L, et al. Neural network predicts need for red blood cell

transfusion for patients with acute gastrointestinal bleeding admitted to the intensive care unit. Sci Rep. 2021;11(1): 8827.

49. Karnad A, Poskitt TR. The automated complete blood cell count: use of the red blood cell volume distribution width and mean platelet volume in evaluating anemia and thrombocytopenia. Arch Intern Med. 1985;145(7):1270–2.
50. May JE, Marques MB, Reddy VVB, Gangaraju R. Three neglected numbers in the CBC: the RDW, MPV, and NRBC count. Cleve Clin J Med. 2019;86:167–72.
51. Stahl D, Lacroix-Desmazes S, Sibrowski W, Kazatchkine MD, Kaveri SV. Red blood cell transfusions are associated with alterations in self-reactive antibody repertoires of plasma IgM and IgG, independent of the presence of a specific immune response toward RBC antigens. Clin Immunol. 2002;105(1): 25–35.
52. Ekber Karabulut A, Çevik Y, Emektar E, Kerem Çorbacioğlu Ş, Dağar S, Yardim O. Analysis of mean platelet volume and red blood cell distribution width in recurrent epistaxis. Turkish J Emerg Med. 2018 Jun 1;18(2):67–70.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Li N, Arnold DM, Down DG, Barty R, Blake J, Chiang F, et al. From demand forecasting to inventory ordering decisions for red blood cells through integrating machine learning, statistical modeling, and inventory optimization. Transfusion. 2022;62:87–99. https://doi.org/10.1111/trf.16739