A risk scoring system integrating postoperative factors for predicting early mortality after major non-cardiac surgery

Dongwoo Chae[1] | Na Young Kim[2] | Hyun Joo Kim[2] | Tae Lim Kim[2] | Su Jeong Kang[2] | So Yeon Kim[2]

[1]Department of Pharmacology, Yonsei University College of Medicine, Seoul, Korea

[2]Department of Anesthesiology and Pain Medicine, Anesthesia and Pain Research Institute, Yonsei University College of Medicine, Seoul, Korea

Correspondence
So Yeon Kim, Department of Anesthesiology and Pain Medicine and Anesthesia and Pain Research Institute, Yonsei University College of Medicine, 50-1 Yonsei-ro, Seodaemun-gu, Seoul 03722, Korea.
Email: kimsy326@yuhs.ac

Dongwoo Chae, Department of Pharmacology, Yonsei University College of Medicine, 50-1 Yonsei-ro, Seodaemun-gu, Seoul 03722, Korea.
Email: dongy@yuhs.ac

Abstract

We aimed to develop a risk scoring system for 1-week and 1-month mortality after major non-cardiac surgery, and assess the impact of postoperative factors on 1-week and 1-month mortality using machine learning algorithms. We retrospectively reviewed the medical records of 21,510 patients who were transfused with red blood cells during non-cardiac surgery and collected pre-, intra-, and postoperative features. We derived two patient cohorts to predict 1-week and 1-month mortality and randomly split each of them into training and test cohorts at a ratio of 8:2. All the modeling steps were carried out solely based on the training cohorts, whereas the test cohorts were reserved for the evaluation of predictive performance. Incorporation of postoperative information demonstrated no significant benefit in predicting 1-week mortality but led to substantial improvement in predicting 1-month mortality. Risk scores predicting 1-week and 1-month mortality were associated with area under receiver operating characteristic curves of 84.58% and 90.66%, respectively. Brain surgery, amount of intraoperative red blood cell transfusion, preoperative platelet count, preoperative serum albumin, and American Society of Anesthesiologists physical status were included in the risk score predicting 1-week mortality. Postoperative day (POD) 5 (neutrophil count × mean platelet volume) to (lymphocyte count × platelet count) ratio, preoperative and POD 5 serum albumin, and occurrence of acute kidney injury were included in the risk score predicting 1-month mortality. Our scoring system advocates the importance of postoperative complete blood count differential and serum albumin to better predict mortality beyond the first week post-surgery.

Study Highlights

WHAT IS THE CURRENT KNOWLEDGE ON THE TOPIC?
Postoperative mortality accounts for 7.7% of all-cause mortality and is the third greatest contributor to death. Hence, risk stratification and prediction of

---

Dongwoo Chae and Na Young Kim contributed equally to this work.

postoperative mortality are crucial for improving global health. Previous studies identified several pre- and intraoperative risk factors for predicting postoperative mortality but largely overlooked postoperative factors.

WHAT QUESTION DID THIS STUDY ADDRESS?

Would incorporation of early postoperative information help improve mortality prediction? If so, what are the most important factors? Are risk factors of 1-week and 1-month mortality different?

WHAT DOES THIS STUDY ADD TO OUR KNOWLEDGE?

A simple risk scoring system predicting 1-week and 1-month mortality after major non-cardiac surgery was developed. Different sets of risk factors were associated with 1-week and 1-month mortality. Incorporation of postoperative information demonstrated no significant benefit in predicting 1-week mortality but led to substantial improvement in predicting 1-month mortality. Routine laboratory tests such as complete blood count (CBC) differentials and serum albumin performed on postoperative day 5 and occurrence of acute kidney injury (AKI) offered crucial information to predict 1-month mortality.

HOW MIGHT THIS CHANGE CLINICAL PHARMACOLOGY OR TRANSLATIONAL SCIENCE?

To efficiently predict postoperative mortality, clinicians should not only consider pre- and intraoperative factors but also postoperative variables. In particular, serum albumin levels and CBC indices can be utilized to predict mortality beyond the first week of surgery. AKI also significantly increases the risk of patient mortality, suggesting the need for effective prevention strategies.

INTRODUCTION

Postoperative death—a global burden accounting for 7.7% of all-cause mortality—is the third greatest contributor to death, following ischemic heart disease and stroke.[1] Globally, more than 200 million patients undergo major non-cardiac surgery each year and 4.2 million of them die within 30 days after surgery.[1,2] Therefore, risk stratification and prediction of postoperative mortality are essential for improving global health. There are several risk scores and models for predicting mortality after major surgery, most of which utilize pre- and intraoperative factors.[3–7] However, postoperative factors may significantly impact postoperative mortality.[8] Moreover, mortality risk factors may change in different postoperative periods.[5]

Neutrophils, lymphocytes, and platelets play essential roles in mediating innate and adaptive immune responses. Systemic inflammation is associated with alterations in the counts of these blood cells, and it often manifests as neutrophilia, lymphopenia, and thrombocytopenia.[9–11] Several studies have revealed significant associations between such alterations and postoperative mortality.[12–16] Moreover, albumin is a negative acute phase reactant that is downregulated during inflammation, irrespective of the patient's nutritional state.[17,18] Hence, pre- and postoperative hypoalbuminemia reportedly increase postoperative

mortality risk.[19–22] Therefore, the evidence suggests that alterations of complete blood count (CBC) and serum albumin levels offer significant prognostic information.

In this study, we aimed to explore the most informative features predictive of early mortality after major non-cardiac surgery using a large pool of candidate features collected during pre-, intra-, and postoperative periods. We used various machine learning algorithms to identify robust predictors of postoperative mortality. In particular, we investigated the added benefit of incorporating postoperative features by comparing prediction models trained using pre- and intraoperative variables with those additionally incorporating postoperative variables. Among postoperative features, we focused on evaluating the impact of time-varying variables, such as CBC with differential and serum albumin. Finally, a simple risk scoring system of 1-week and 1-month mortality after surgery was developed with the selected features.

METHODS

Patients

We retrospectively collected data from electronic medical records of a university hospital that included 21,510 patients, with an American Society of Anesthesiologists

(ASA) physical status of I–IV, who were transfused with at least one unit of red blood cells (RBCs) during non-cardiac surgery between November 2005 and January 2020. We only included the patients receiving at least one unit of RBCs during surgery due to the higher risk of early mortality in these patients,[23–26] thus filtering out minor surgeries associated with negligible mortality risk. The study was approved by the Institutional Review Board and Hospital Research Ethics Committee of Severance Hospital, Yonsei University Health System, Seoul, Korea (number: 4-2020-0590). All information was anonymized, and the requirement for informed consent was waived due to the retrospective study design. All methods were carried out in accordance with relevant guidelines and regulations.

Derivation of analysis datasets

To identify risk factors that change with time, we defined two patient cohorts based on their survivorship. All patients who were alive immediately after surgery were defined as cohort 1, and the mortality event was defined as death within 1 week of surgery. Cohort 2 was derived by excluding patients who died before 1 week, and the events of interest were deaths between 1 week and 1 month of surgery. This scheme yielded 21,510 and 21,374 patients for cohorts 1 and 2, respectively, because 136 patients of cohort 1 died in the first postoperative week. Prediction models were subsequently developed using candidate features that included pre-, intra-, and postoperative

information. Prior to training the models, we randomly split the cohorts into training and test cohorts in a ratio of 8:2. Hereafter, all the modeling steps were carried out solely based on the training cohorts and the test cohorts were reserved for evaluation of predictive performance. All features were z-score-normalized prior to model development. The analysis workflow is schematically illustrated in Figure 1.

Candidate features

The original dataset consisted of the following features: patient characteristics (age, sex, body mass index [BMI], ASA physical status, and comorbidities), pre- and postoperative laboratory parameters (CBC, serum albumin, and estimated glomerular filtration rate [eGFR] calculated using the Chronic Kidney Disease Epidemiology Collaboration equation[27]), intraoperative factors, such as amounts of RBC transfusions, anesthesia method (general vs. spinal, inhalation vs. total intravenous anesthesia [TIVA]), anesthesia duration, blood loss, urine output, emergency surgery, type of surgery and field of surgery, and postoperative acute kidney injury (AKI) occurrence. For preoperative laboratory parameters, the latest measurement within 2 months before surgery was chosen. Postoperative laboratory parameters were collected postoperative day (POD) 0 (immediately after surgery) to 5 days. AKI was defined as an increase in serum creatinine by 50% within 7 days postoperatively, or an increase in serum creatinine by 0.3 mg/dL within 2 days
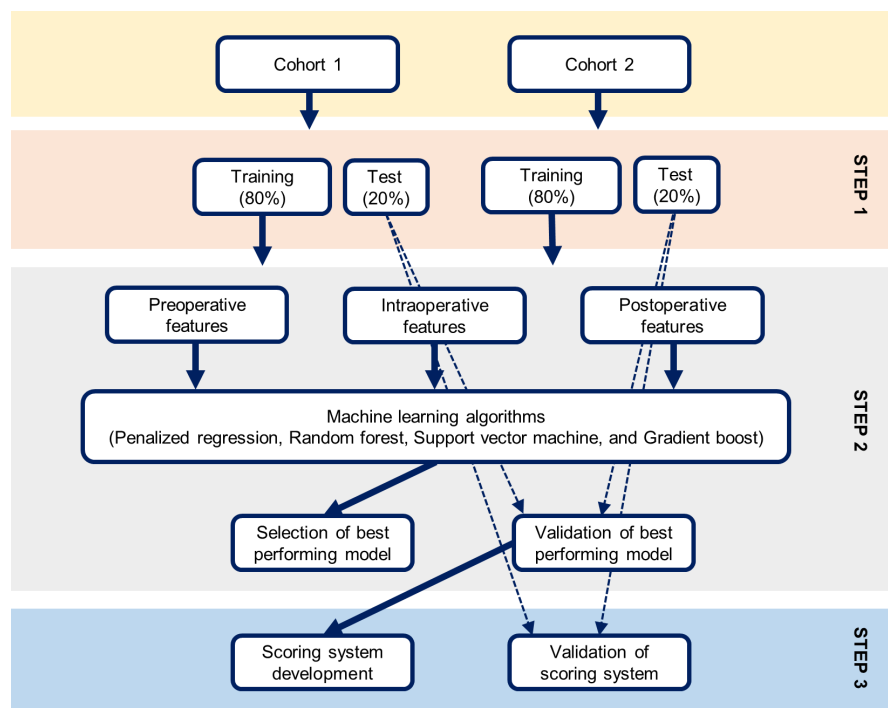
FIGURE 1 Description of the analysis workflow. Two patient cohorts were randomly split into training (80%) and test (20%) cohorts (step 1). The training cohorts were used to train various machine learning models, out of which the best performing model was selected and validated (step 2). The most important features of the validated model were used to develop the risk scoring system, and its performance was assessed using the test cohorts (step 3).

postoperatively, based on the Kidney Disease Improving Global Outcomes Work Group criteria.[28] Among the original features, the following features were used for training the models:

(i) Patient characteristics: age, sex, BMI, ASA physical status, and comorbidities (hypertension, diabetes mellitus, atrial fibrillation, chronic kidney disease, cerebrovascular disease, and coronary artery disease).
(ii) Preoperative features: preoperative CBC (neutrophil count, lymphocyte count, platelet count, mean platelet volume [MPV], and hematocrit), serum albumin, and eGFR.
(iii) Intraoperative features: anesthesia duration, type of anesthesia (TIVA or spinal), blood loss, albumin administration, RBC transfusion, platelet transfusion, fresh frozen plasma transfusion, urine output per hour, emergency surgery, type of surgery (laparoscopic surgery and cancer surgery), and field of surgery.
(iv) Postoperative features: CBC (neutrophil count, lymphocyte count, platelet count, MPV, and hematocrit) and serum albumin measured on PODs 0 to 5 and AKI occurrence.

Missing values were imputed using multiple imputation with chained equations, implemented via the R package "mice."[29]

Model development

Penalized regression (least absolute shrinkage and selection operator [LASSO], Ridge, and Elastic Net), random forest, support vector machine, and gradient boost methods were trained on the two cohorts. Python packages ("scikit-learn" and "lightgbm") were used for implementing each of the abovementioned models. Four-fold cross validation was used to tune the hyperparameters. The best performing model was identified based on cross-validated log loss. Permutation feature importance values were subsequently generated to identify the most essential features based on the training cohorts.

In the initial stage, only pre- and intraoperative features were used for model training. Next, we additionally incorporated postoperative features. For cohort 1, postoperative information was limited to those collected on POD 0. For cohort 2, all information available between PODs 0 and 5 were considered. We first used all available measurements between PODs 0 and 5 to train a reference LASSO model. We next explored whether a single timepoint chosen from PODs 0, 1, 2, 3, 4, or 5 would provide comparable predictive performance. To this end, we additionally trained

LASSO models using measurements taken from one of PODs 0, 1, 2, 3, 4, or 5. If cross-validated log loss of the best performing model was not greater than 5% of the reference log loss, we chose the simpler prediction scheme of using only a single timepoint value for subsequent model development.

Assessment of predictive performance

Predictive performances were assessed on the test cohorts using four different metrics—area under the receiver operating characteristic curve (AUROC), area under the precision-recall curve (AUPR), Brier score, and log loss. Performances of models developed using pre- and intraoperative information were compared with those additionally incorporating postoperative information.

Final clinical score development

Top-ranking features across all cohorts were identified from the feature importance values. Optimal binning algorithm implemented via the "OptBinning" Python package[30] was then used to derive the final risk scoring system. The incremental contribution of each of the risk factors was assessed based on the mean relative percentage increase in log loss upon random permutations, repeated 100 times.

RESULTS

Patient characteristics are summarized in Table 1. Notably, 40.1% of included patients underwent cancer surgery, and of all, abdominal surgery was the most common (19.2%). The mean anesthesia duration and intraoperative blood loss were 5.5 h and 1158 ml, respectively, with 35.7% of patients receiving greater than or equal to three units of intraoperative RBC transfusion.

Exploratory data analysis

The frequencies and percentages of the different causes of 1-week and 1-month mortality after surgery are shown in Table S1. Mortality due to hemorrhagic events was the most frequent in the first week and decreased thereafter. Multiple organ dysfunction syndrome (MODS) was the most frequent cause of death by 2–4 weeks. Respiratory failure was one of the lowest contributors to death in the first week but became the third most frequent cause by 2–4 weeks. Septic shock was a relatively steady contributor

TABLE 1 Patient characteristics ($n = 21{,}510$)

Variable	
Demographics	
Age, year	60.0 (15.6)
Female	12,871 (59.8%)
Body mass index, kg/m^2	23.7 (3.7)
ASA physical status	
I	4625 (21.5%)
II	9850 (45.8%)
III	6206 (28.8%)
IV	829 (3.9%)
Comorbidities	
Hypertension	10,945 (50.9%)
Diabetes mellitus	5376 (25.0%)
Atrial fibrillation	367 (1.71%)
Chronic kidney disease	432 (2.0%)
Cerebrovascular disease	1510 (7.0%)
Coronary artery disease	629 (2.9%)
Preoperative laboratory values	
Neutrophil count, $\times 10^3/\mu L$	5.00 (3.45)
Lymphocyte count, $\times 10^3/\mu L$	1.74 (0.81)
Platelet count, $\times 10^3/\mu L$	259 (101)
Mean platelet volume, fL	8.7 (1.2)
Hematocrit (%)	36.3 (5.7)
Serum albumin, g/dL	4.0 (0.6)
eGFR, ml/min/1.73 m^2	87.9 (23.1)
Intraoperative factors	
TIVA	1759 (8.2%)
Duration of anesthesia, h	5.5 (3.3)
Blood loss, ml	1158 (1349)
Urine output, ml/h	135 (112)
Red blood cell transfusion	
1/2/≥3 units	8058 (37.5%)/5769 (26.8%)/7683 (35.7%)
Emergency surgery	2846 (13.2%)
Type of surgery	
Laparoscopic surgery	1158 (5.4%)
Cancer surgery	8628 (40.1%)
Field of surgery	
Abdominal	4121 (19.2%)
Thoracic	545 (2.5%)
Brain	2771 (12.9%)
Head and neck	1352 (6.3%)
Breast	107 (0.5%)
Obstetric/gynecological	310 (1.4%)/972 (4.5%)
Kidney	575 (2.7%)

TABLE 1 (Continued)

Variable	
Prostate/bladder and urinary	643 (3.0%)/684 (3.2%)
Cervical/thoraco-lumbar spine	440 (2.0%)/4032 (18.7%)
Upper/lower extremities	185 (0.9%)/3534 (16.4%)
Hip	1091 (5.1%)
Skin, soft tissue	60 (0.3%)
Postoperative AKI	1779 (8.3%)

Notes: Values are expressed as mean (standard deviation) or number of patients (percentage).

Abbreviations: ASA, American Society of Anesthesiologists; AKI, acute kidney injury; eGFR, estimated glomerular filtration rate; TIVA, total intravenous anesthesia.

to death and hypovolemic and cardiogenic shocks were minor contributors.

Training machine learning models

We trained various machine learning models on the two training cohorts using pre- and intraoperative features. Hyper-parameter tuning and model selection were carried out based on comparison of cross-validated log loss. LightGBM and ElasticNet yielded the best prediction performances for cohorts 1 and 2, respectively (Table S2).

Next, we trained machine learning models on the two training cohorts using additional information collected in the postoperative period. For cohort 1, routinely performed laboratory values on POD 0 were used. For cohort 2, occurrence of AKI and all laboratory values measured between PODs 0 and 5 were initially used to train a reference LASSO model. Next, LASSO models were retrained by selecting only a single timepoint from PODs 0 to 5. The cross-validated log loss of the model using all available information was 0.0560. The best performing single-timepoint model was shown to yield a cross-validated log loss of 0.0568 based on measurements taken on POD 5. Because the log loss was not greater than 5% of the reference log loss, we concluded that POD 5 offers the optimal timepoint for subsequent model building, and, regardless of the machine learning algorithm used, additional incorporation of postoperative variables improved prediction performance in cohort 2, but not in cohort 1. Hence, we only incorporated postoperative information for cohort 2. LightGBM was the best performing model for both cohorts (Table S2).
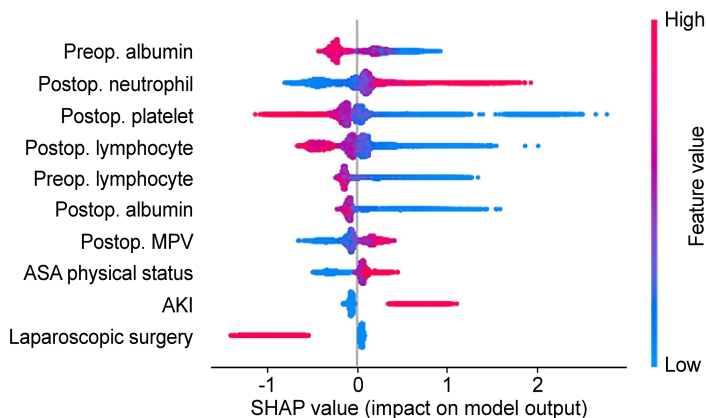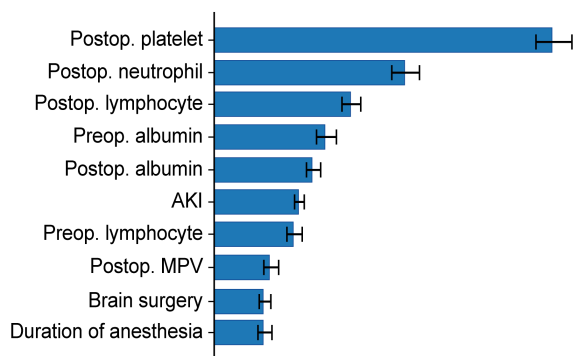
Feature importance assessment

The permutation feature importance plots of the top 10 features and SHapley Additive exPlanations (SHAP)
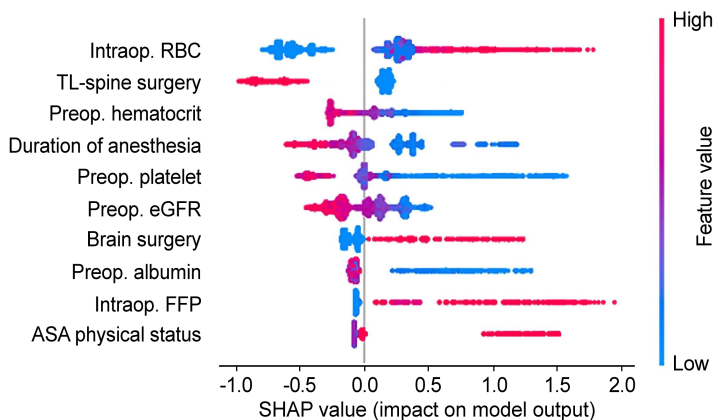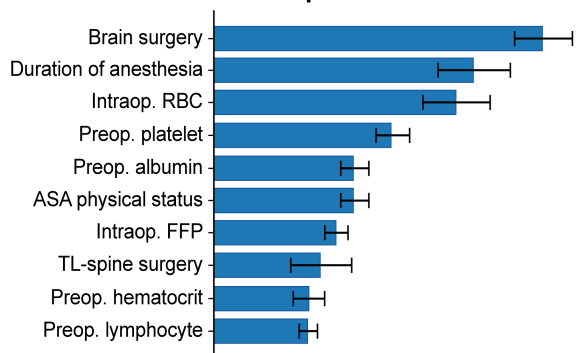
FIGURE 2 The feature importance and SHapley Additive exPlanations (SHAP) summary plots were generated based on the best performing lightGBM models. AKI, acute kidney injury; ASA, American Society of Anesthesiologists; eGFR, estimated glomerular filtration rate; FFP, fresh frozen plasma; Intraop, intraoperative; MPV, mean platelet volume; Preop, preoperative; Postop, postoperative; RBC, red blood cell; TL, thoracolumbar.

summary plots of lightGBM model are shown in Figure 2, evaluated using the training cohorts. In cohort 1, brain surgery, duration of anesthesia, and amount of intraoperative RBC administration were the top three factors, followed by preoperative platelet and albumin levels. In cohort 2, the highest-ranking features were POD 5 platelet count, neutrophil count, and lymphocyte count, followed by preoperative albumin level, POD 5 albumin level, and AKI occurrence. POD 5 MPV was found as the eighth most important feature. Hence, CBC differentials arose as crucial predictors of 1-month mortality. The SHAP summary plot revealed that lower platelet count, higher neutrophil count, lower lymphocyte count, and higher MPV were associated with a higher mortality risk. Interestingly, duration of anesthesia was inversely correlated with 1-week mortality risk based on the SHAP summary plot. Closer scrutiny revealed a biphasic pattern between duration of anesthesia and incidence of mortality in cohort 1 such that durations of less than 1, 1–2, 2–3, 3–4, 4–5, and greater than or equal to 5 h were associated with 0%, 0.59%, 0.91%, 1.19%, 0.53%, and 0.44% mortality, respectively. For

mortality in cohort 2, a longer duration of anesthesia was positively correlated with mortality.

Evaluation of predictive performances

The predictive performances of the best performing lightGBM models were evaluated using log loss, AUROC, Brier score, and AUPR using the test cohorts (Table 2). As was suggested from comparison of cross-validated log-loss, incorporation of postoperative features led to a substantial improvement in all four performance metrics in the test dataset of cohort 2, with nearly a twofold increase in AUPR.

Feature selection and engineering for risk score development

For cohort 1, we selected brain surgery, the amount of intraoperative RBC transfusion, preoperative platelet count, preoperative albumin level, and ASA physical status as the

Cohort	Features	Log loss	AUROC	Brier score	AUPR
1	Pre- and intraoperative	0.0307	87.76%	0.00621	8.65%
	Pre-, intra-, and postoperative	0.0302	89.17%	0.00610	10.82%
2	Pre- and intraoperative	0.0728	86.32%	0.0166	17.75%
	Pre-, intra-, and postoperative	0.0619	90.53%	0.0146	35.48%

TABLE 2 Performance comparisons of lightGBM model evaluated using the test datasets

Abbreviations: AUROC, area under the receiver operating characteristic curve; AUPR, area under precision-recall curve.

features to be included in the final risk score. For cohort 2, we explored effective methods of feature engineering. Given the signature pattern of neutrophilia, lymphocytopenia, higher MPV, and thrombocytopenia being associated with higher mortality risk, we tested neutrophil-to-lymphocyte ratio (NLR) and MPV-to-platelet ratio (MPR) as substitute features. Given the highly right-skewed distribution of both NLR and MPR values, we applied logarithmic transformation with base two to both. Training lightGBM models by substituting the four CBC-related features with $\log_2$ NLR and $\log_2$ MPR led to cross-validated log loss of 0.0554. We also created a new composite index, (neutrophil count × mean platelet volume) to (lymphocyte count × platelet count) ratio (NMLPR), by multiplying NLR by MPR. The cross-validated log loss of the lightGBM model using $\log_2$ NMLPR was 0.0558, which was not greater than 5% of the cross-validated log loss of 0.0553 in the lightGBM model using the original CBC differentials.

In cohort 2, comparison of various performance metrics using the test dataset additionally confirmed that use of NMLPR, despite lumping four variables into a single number, was not substantially inferior (Table S3). We thus concluded that NMLPR constitutes an effective mortality risk indicator with high data compression efficiency and used it as a variable representing the overall effect of CBC alterations during subsequent risk score development. The other selected risk factors were preoperative albumin, POD 5 albumin, and AKI.

Final risk score development

An optimal binning algorithm was used to derive the final risk scoring system. The scoring scheme and the predictive performance evaluated on the test datasets are shown in Table 3. The total score can be converted to a probability estimate based on the following equation:

$$\text{Probability of patient death} = \frac{1}{1 + \exp\left(-0.1 \times \text{total score}\right)}$$

The log loss, AUROC, Brier score, and AUPR of the scoring system assessed using the two test cohorts showed that log loss and Brier score were slightly inferior to those of the lightGBM models. Nevertheless, the performance was

comparable in terms of AUROC and AUPR (Tables 2 and 3). Random permutation of each of the risk factors showed that the amount of intraoperative RBC transfusion, brain surgery, preoperative platelet count, and preoperative albumin showed a similar incremental contribution of 9–12% in predicting 1-week mortality. On the contrary, POD 5 NMLPR showed a dominant incremental contribution to the prediction of 1-month mortality with a 43% relative increase in log loss upon random permutation (Table S4). To verify how well predicted probabilities matched observed probabilities, predictions were binned into intervals of [0, 0.025], [0.025, 0.05], [0.05, 0.1], and [0.1, 1], yielding a similar number of mortality events across all bins, and observed mortality rates (i.e., number of deaths/total number of patients) within each interval were visualized (Figure 3). The results showed that the actual mortality rates in the corresponding intervals were 0.00, 0.03, 0.08, and 0.14 for cohort 1 and 0.01, 0.04, 0.09, and 0.25 for cohort 2, respectively.

DISCUSSION

We developed a risk scoring system of 1-week and 1-month mortality after major non-cardiac surgery. This study addressed two core issues: (1) the impact of postoperative factors on early postoperative mortality, and (2) the change of risk factor profile with postoperative time. Postoperative information did not offer much benefit for predicting 1-week mortality when considered in addition to pre- and intraoperative information. This was in stark contrast to 1-month mortality prediction where postoperative information—including serum albumin and CBC indices—were crucial in improving the prediction accuracy. Despite the parsimonious inclusion of predictor variables, the final scoring system demonstrated good predictive performance (i.e., AUROC >80%) in both 1-week and 1-month mortality.

Several risk stratification tools for predicting mortality after non-cardiac surgery have been proposed, most of which utilized pre- and/or intraoperative factors as predictors.[3–7,31] For example, Physiological and Operative Severity Score for the enUmeration of Mortality and morbidity (POSSUM) and Portsmouth POSSUM scores—the most frequently used tools in heterogeneous surgical

TABLE 3 Final risk scoring system

| Endpoint/performance	Feature	Range	Points
1-week mortality Log loss: 0.0330 AUROC: 84.58% Brier score: 0.00625 AUPR: 9.31%	Brain surgery	No	−4
		Yes	+13
	Amount of intraoperative RBC transfusion (pack)	<2	−9
		2–4	−1
		4–10	+1
		≥10	+14
	Preoperative platelet count (cell/nL)	<120	+13
		120–200	0
		200–220	−16
		≥220	−3
	Preoperative albumin (g/dL)	<3.1	+11
		3.1–3.7	0
		3.7–4.6	−4
		≥4.6	−11
	ASA physical status	1–2	−5
		3–4	+5
1-month mortality Log-loss: 0.0640 AUROC: 90.66% Brier score: 0.0155 AUPR: 25.91%	POD 5 NMLPR ($\times 10^{-8}$ nL^2/cell)	<100	−22
		100–300	−8
		300–1000	+4
		≥1000	+19
	Preoperative albumin (g/dL)	<3.1	+8
		3.1–3.8	+2
		3.8–4.1	0
		≥4.1	−7
	POD 5 albumin (g/dL)	<2.7	+6
		2.7–3.0	+1
		3.0–3.3	−2
		≥3.3	−5
	AKI	No	−2
		Yes	+8

Abbreviations: AKI, acute kidney injury; ASA, American Society of Anesthesiologists; AUROC, area under the receiver operating characteristic curve; AUPR, area under precision-recall curve; NMLPR, (neutrophil count × mean platelet volume) to (lymphocyte count × platelet count) ratio; POD, postoperative day; RBC, red blood cell.

patients—included 12 physiological and six operative factors.[31] However, these models had a core limitation; the impact of postoperative factors was ignored. Because patient mortality results from various events cannot be effectively predicted preoperatively, a realistic prediction scheme should allow for the possibility of updating of the risk estimates with newly available information. Our study is one of the few to investigate the effects of postoperative factors in determining mortality risk,[8] parting from a rigid framework wherein only pre- and/or intraoperative information was used. Our approach uses newly acquired postoperative information to reassess mortality risk at 1 week after surgery. Such dynamic evaluation involved separately

identifying the most significant risk factors of 1-week and 1-month mortality. To this end, we defined two patient cohorts by censoring mortality events at different timepoints, thereby artificially mimicking clinical trials with different start and end times. By elucidating significant mortality risk factors in the two cohorts, we aimed to gain a simple and clear picture of what risk factors are important to predict 1-week and 1-month mortality.

The identified risk factors for the two periods—day 0 to 1 week and 1 week to 1 month— seem to reflect the major causes of death in the corresponding periods. In line with the observation of intracranial hemorrhage and hemorrhagic shock constituting the first and third most
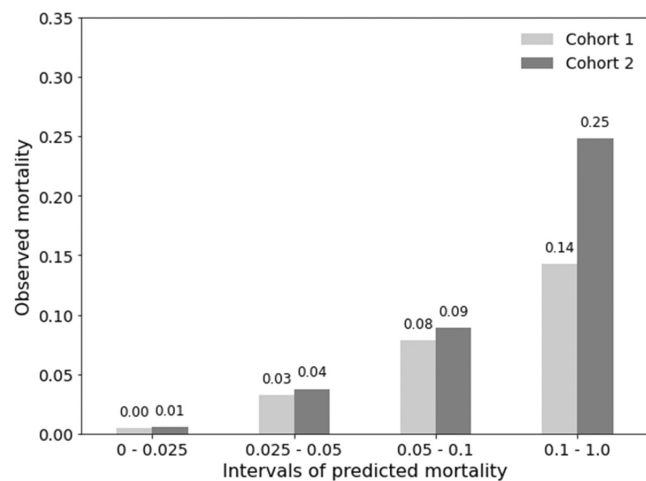
FIGURE 3 The calibration plot of the final risk scoring system developed based on the most important features showing observed versus predicted probabilities of mortality.

frequent causes of death in the first week (Table S1), brain surgery, the amount of RBC transfusion, and preoperative platelet count were identified as the top risk factors. Meanwhile, the increasing incidences of respiratory failure (often implying pneumonia), septic shock, and MODS between 2 and 4 weeks suggest that markers of systemic inflammation and/or infection would become more important beyond the first week of surgery. The identification of CBC differentials and serum albumin as the most crucial factors support this possibility because these are widely known biomarkers of inflammation and sepsis.

We included a novel CBC-based index, NMLPR, in the final clinical scoring system, which strongly correlated with postoperative mortality risk. NMLPR is a composite index of NLR×MPR, both of which are known predictors of postoperative and other disease-related mortalities.[12,14–16,32–34] The use of composite indices offers two main advantages: (1) they compress information and (2) increase the diagnostic and prognostic specificity of individual biomarkers. With systemic inflammation, the lymphocyte count decreases and the neutrophil count increases.[9] NLR, which is a composite index of neutrophils and lymphocytes, reflects changes in both cell types and is therefore a more sensitive marker of systemic inflammation than measures of either cell type alone.[35] Platelets are increasingly recognized as important contributors to inflammation and thrombocytopenia is a common finding in patients with critical illness.[36] MPR, which is a composite index of MPV and platelet count, was shown to predict poor clinical outcome in patients admitted to the surgical intensive care unit.[37] In addition, a high MPR was an independent predictor of 28-day mortality in patients with severe sepsis.[38] Because systemic inflammation is responsible for increasing both NLR and MPR, NMLPR might be a more robust indicator of inflammatory status than either NLR or MPR. Indeed, our analysis showed that NMLPR was as effective as NLR and MPR when used for predicting 30-day mortality.

The close association of pre- and postoperative serum albumin level with postoperative mortality is well-documented,[18–21] and our study strengthens this view. The definition of hypoalbuminemia differs from study to study and ranges from less than 2.5 g/dL to less than 4.0 g/dL.[19–21] The optimal binning algorithm used to construct our risk scoring system suggested a binning scheme of dividing albumin ranges into less than 3.1, 3.1–3.8, 3.8–4.1, and greater than or equal to 4.1 g/dL in cohort 1. However, a different binning scheme was chosen for cohort 2. It seems unlikely that there exists a sharp threshold demarcating increased risk from no risk. Rather, a decline in the albumin level seems to result in graded increases in mortality risk. Whether lower albumin is a causative factor of increased mortality risk cannot be answered from a retrospective analysis. Further studies investigating the benefit of increasing albumin level during perioperative periods may be needed; according to a previous study, human albumin administration in the early postoperative periods was not beneficial in correcting hypoalbuminemia or improving clinical outcomes following gastrointestinal surgery.[39]

Preoperative hematocrit was found as the ninth most important feature in cohort 1, whereas pre- or postoperative hematocrit was not selected among the top 10 features in cohort 2. As a result, hematocrit was not included in our scoring system, although preoperative anemia reportedly increases mortality after non-cardiac surgery.[4,40,41] Intraoperative RBC transfusion of 1 or 2 units is known to be associated with a higher risk of 30-day mortality and morbidity after non-cardiac surgery.[24–26] In our study, all patients received an intraoperative RBC transfusion and there was a definite association between the amount of packed RBC and the risk of 1-week mortality, with a dramatic increased mortality risk when transfused with greater than or equal to 10 units. However, intraoperative RBC transfusion was only associated with 1-week mortality but not 1-month mortality, which implies that RBC transfusion is no longer a main determinant of mortality after 1 week of non-cardiac surgery.

The ASA physical status is the most widely used tool to classify patient preoperative health, and thus has been incorporated into several risk scores for postoperative mortality.[42] Similarly, ASA physical status was selected in our final scoring system for 1-week mortality. However, the importance of ASA physical status on mortality was trivial after 1 week of surgery. AKI occurrence was found as a significant risk factor increasing 1-month mortality, which is in concordance with previous studies demonstrating an increased hospital mortality risk in patients with than without postoperative AKI.[43,44] Nevertheless, its impact on mortality compared to CBC and albumin level seems less pronounced.

The predictive performances of the final risk score show overall better performances than those found in the literature. Based on a recent systematic review, previous models using preoperative variables have reported AUROC ranging from 82% to 93%, excluding one study conducted in 1998 reporting AUROC of 0.68.[7] Due to the heterogeneity of the underlying patient population used in the different studies, however, a direct comparison of predictive performances is not always possible. To assess the impact of the year the surgery was performed, we compared the predictive performances of our scoring system across different time periods ranging from 2005 to 2020. No specific trend was noted, and the associated variability appeared largely random (Table S5).

Our study has some limitations. First, data were retrospectively collected. Hence, no clear conclusion can be drawn regarding whether high NMLPR and hypoalbuminemia are true predictors or mere epiphenomena of high mortality. This warrants further research to clarify the clinical implications and physiology underlying the possible role of the above-mentioned parameters in postoperative mortality prediction for definite clinical acceptance. Second, the patient population might not be representative of the general surgical population. We used data of patients who underwent non-cardiac surgery requiring intraoperative RBC transfusion in one of the largest tertiary hospitals in South Korea. Hence, the predicted mortality probability of our model should not be extrapolated directly to all postoperative patients, except after a careful baseline mortality hazard recalibration. Despite the abovementioned limitations, our study population was sufficiently large, and therefore the study was satisfactorily powered to estimate the smallest effect sizes.

In conclusion, we developed a risk scoring system that integrated pre-, intra-, and postoperative factors to predict 1-week and 1-month mortality in patients undergoing major non-cardiac surgery with RBC transfusion. Albumin and CBC differentials were revealed as main factors and NMLPR is proposed as a novel and efficient CBC-based composite index for mortality prediction. Postoperative laboratory values were found to constitute key information to predict 1-month mortality but not 1-week mortality. This result clearly suggests the time-varying nature of postoperative mortality risk and alarms the physicians to sensitively attend to the patients' changing status. The dynamic framework that we used may foster similar approaches in future studies to effectively assess time-varying postoperative mortality predictors.

AUTHOR CONTRIBUTIONS

D.C., N.Y.K., and S.Y.K. wrote the manuscript. D.C., N.Y.K., and S.Y.K. designed the research. D.C., N.Y.K., S.Y.K., H.J.K., T.L.K., and S.J.K. performed the research. D.C., N.Y.K., H.J.K., and S.Y.K. analyzed the data.

CONFLICT OF INTEREST

ORCID

Na Young Kim https://orcid.org/0000-0003-3685-2005
So Yeon Kim https://orcid.org/0000-0001-5352-157X

REFERENCES

1. Nepogodiev D, Martin J, Biccard B, et al. Global burden of postoperative death. *Lancet*. 2019;393(10170):401.
2. Weiser TG, Regenbogen SE, Thompson KD, et al. An estimation of the global volume of surgery: a modelling strategy based on available data. *Lancet*. 2008;372(9633):139-144.
3. Moonesinghe SR, Mythen MG, Das P, Rowan KM, Grocott MP. Risk stratification tools for predicting morbidity and mortality in adult patients undergoing major surgery: qualitative systematic review. *Anesthesiology*. 2013;119(4):959-981.
4. Chan DXH, Sim YE, Chan YH, Poopalalingam R, Abdullah HR. Development of the Combined Assessment of Risk Encountered in Surgery (CARES) surgical risk calculator for prediction of postsurgical mortality and need for intensive care unit admission risk: a single-center retrospective study. *BMJ Open*. 2018;8(3):e019427.
5. Mahmud N, Fricker Z, Hubbard RA, et al. Risk prediction models for post-operative mortality in patients with cirrhosis. *Hepatology*. 2021;73(1):204-218.
6. Fritz BA, Cui Z, Zhang M, et al. Deep-learning model for predicting 30-day postoperative mortality. *Br J Anaesth*. 2019;123(5):688-695.
7. Reilly JR, Gabbe BJ, Brown WA, Hodgson CL, Myles PS. Systematic review of perioperative mortality risk prediction models for adults undergoing inpatient non-cardiac surgery. *ANZ J Surg*. 2021;91(5):860-870.
8. Blanco JF, da Casa C, Pablos-Hernández C, González-Ramírez A, Julián-Enríquez JM, Díaz-Álvarez A. 30-day mortality after hip fracture surgery: influence of postoperative factors. *PLoS One*. 2021;16(2):e0246963.
9. Pietras EM. Inflammation: a key regulator of hematopoietic stem cell fate in health and disease. *Blood*. 2017;130(15):1693-1698.
10. Kim SJ, Davis RP, Jenne CN. Platelets as modulators of inflammation. *Semin Thromb Hemost*. 2018;44(2):91-101.
11. Guthrie GJ, Charles KA, Roxburgh CS, Horgan PG, McMillan DC, Clarke SJ. The systemic inflammation-based neutrophil-lymphocyte ratio: experience in patients with cancer. *Crit Rev Oncol Hematol*. 2013;88(1):218-230.
12. Chae YJ, Lee J, Park JH, Han DG, Ha E, Yi IK. Late mortality prediction of neutrophil-to-lymphocyte and platelet ratio in patients with trauma who underwent emergency surgery: a retrospective study. *J Surg Res*. 2021;267:755-761.
13. Kertai MD, Zhou S, Karhausen JA, et al. Platelet counts, acute kidney injury, and mortality after coronary artery bypass grafting surgery. *Anesthesiology*. 2016;124(2):339-352.
14. Wang Q, Li J, Wang X. The neutrophil-lymphocyte ratio is associated with postoperative mortality of cardiac surgery. *J Thorac Dis*. 2021;13(1):67-75.
15. Tan TP, Arekapudi A, Metha J, Prasad A, Venkatraghavan L. Neutrophil-lymphocyte ratio as predictor of mortality and morbidity in cardiovascular surgery: a systematic review. *ANZ J Surg*. 2015;85(6):414-419.

16. Kim NY, Chun D-H, Kim SY, et al. Prognostic value of systemic inflammatory indices, NLR, PLR, and MPV, for predicting 1-year survival of patients undergoing cytoreductive surgery with HIPEC. *J Clin Med*. 2019;8(5):589.

17. Ishida S, Hashimoto I, Seike T, Abe Y, Nakaya Y, Nakanishi H. Serum albumin levels correlate with inflammation rather than nutrition supply in burns patients: a retrospective study. *J Med Investig*. 2014;61(3.4):361-368.

18. Ahmed MS, Jadhav AB, Hassan A, Meng QH. Acute phase reactants as novel predictors of cardiovascular disease. *ISRN Inflamm* 2012;2012:953461, 1, 18.

19. Karas PL, Goh SL, Dhital K. Is low serum albumin associated with postoperative complications in patients undergoing cardiac surgery? *Interact Cardiovasc Thorac Surg*. 2015;21(6):777-786.

20. Larsen PB, Liest S, Hannani D, Jørgensen HL, Sørensen LT, Jørgensen LN. Preoperative hypoalbuminemia predicts early mortality following open abdominal surgery in patients above 60 years of age. *Scand J Surg*. 2019;110(1):29-36.

21. Truong A, Hanna MH, Moghadamyeghaneh Z, Stamos MJ. Implications of preoperative hypoalbuminemia in colorectal surgery. *World J Gastrointestinal Surg*. 2016;8(5):353-362.

22. Labgaa I, Joliat G-R, Kefleyesus A, et al. Is postoperative decrease of serum albumin an early predictor of complications after major abdominal surgery? A prospective cohort study in a European centre. *BMJ Open*. 2017;7(4):e013966.

23. Marik PE, Corwin HL. Efficacy of red blood cell transfusion in the critically ill: a systematic review of the literature. *Crit Care Med*. 2008;36(9):2667-2674.

24. Bernard AC, Davenport DL, Chang PK, Vaughan TB, Zwischenberger JB. Intraoperative transfusion of 1 U to 2 U packed red blood cells is associated with increased 30-day mortality, surgical-site infection, pneumonia, and sepsis in general surgery patients. *J Am Coll Surg*. 2009;208(5):931-937. e931-932; discussion 938-939.

25. Glance LG, Dick AW, Mukamel DB, et al. Association between intraoperative blood transfusion and mortality and morbidity in patients undergoing noncardiac surgery. *Anesthesiology*. 2011;114(2):283-292.

26. Karkouti K, Stukel TA, Beattie WS, et al. Relationship of erythrocyte transfusion with short- and long-term mortality in a population-based surgical cohort. *Anesthesiology*. 2012;117(6):1175-1183.

27. Rule AD. The CKD-EPI equation for estimating GFR from serum creatinine: real improvement or more of the same? *Clin J Am Soc Nephrol*. 2010;5(6):951-953.

28. Khwaja A. KDIGO clinical practice guidelines for acute kidney injury. *Nephron Clin Pract*. 2012;120(4):c179-c184.

29. van Buuren S, Groothuis-Oudshoorn K. mice: multivariate imputation by chained equations in R. *J Stat Softw*. 2011;45(3):1-67.

30. Navas-Palencia G. Optimal binning: mathematical programming formulation. *arXiv*. 2001;arXiv:2001.08025 [cs.LG].

31. Prytherch DR, Whiteley MS, Higgins B, Weaver PC, Prout WG, Powell SJ. POSSUM and Portsmouth POSSUM for predicting mortality. *Br J Surg*. 1998;85(9):1217-1220.

32. Liu C, Zhou Y, He X, et al. Mean platelet volume/platelet count ratio predicts long-term mortality in patients with infective endocarditis. *Biomark Med*. 2020;14(4):293-302.

33. Cho J, Lee S, Uh Y, Lee J-H. Usefulness of mean platelet volume to platelet count ratio for predicting the risk of mortality in community-acquired pneumonia. *Arch Med Sci*. 2020;16(6):1327-1335.

34. Li J, Li Y, Sheng X, et al. Combination of mean platelet volume/platelet count ratio and the APACHE II score better predicts the short-term outcome in patients with acute kidney injury receiving continuous renal replacement therapy. *Kidney Blood Press Res*. 2018;43(2):479-489.

35. Zahorec R. Ratio of neutrophil to lymphocyte counts--rapid and simple parameter of systemic inflammation and stress in critically ill. *Bratisl Lek Listy*. 2001;102(1):5-14.

36. Levi M. Platelets in Critical Illness. *Semin Thromb Hemost*. 2016;42(3):252-257.

37. Djordjevic D, Rondovic G, Surbatovic M, et al. Neutrophil-to-lymphocyte ratio, monocyte-to-lymphocyte ratio, platelet-to-lymphocyte ratio, and mean platelet volume-to-platelet count ratio as biomarkers in critically ill and injured patients: which ratio to choose to predict outcome and nature of bacteremia? *Mediat Inflamm*. 2018;2018:3758068-3758015.

38. Oh GH, Chung SP, Park YS, et al. Mean platelet volume to platelet count ratio as a promising predictor of early mortality in severe sepsis. *Shock*. 2017;47(3):323-330.

39. Yuan XY, Zhang CH, He YL, et al. Is albumin administration beneficial in early stage of postoperative hypoalbuminemia following gastrointestinal surgery? a prospective randomized controlled trial. *Am J Surg*. 2008;196(5):751-755.

40. Musallam KM, Tamim HM, Richards T, et al. Preoperative anaemia and postoperative outcomes in non-cardiac surgery: a retrospective cohort study. *Lancet*. 2011;378(9800):1396-1407.

41. Liew LQ, Teo WW, Seet E, et al. Factors predicting one-year post-surgical mortality amongst older Asian patients undergoing moderate to major non-cardiac surgery – a retrospective cohort study. *BMC Surg*. 2020;20(1):11.

42. Mayhew D, Mendonca V, Murthy BVS. A review of ASA physical status – historical perspectives and modern developments. *Anaesthesia*. 2019;74(3):373-379.

43. Hobson C, Ozrazgat-Baslanti T, Kuxhausen A, et al. Cost and mortality associated with postoperative acute kidney injury. *Ann Surg*. 2015;261(6):1207-1214.

44. Chaudery H, MacDonald N, Ahmad T, et al. Acute kidney injury and risk of death after elective surgery: prospective analysis of data from an international cohort study. *Anesth Analg*. 2019;128(5):1022-1029.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.