

Limitations of risk stratification in CABG: Calibration across elective and emergency cohorts at a regional center

✉ Anıl Akbaş, ✉ Tamer Cebe, ✉ Ahmet İbrahim Balkaya

Clinic of Cardiovascular Surgery, Batman Training and Research Hospital, Batman, Türkiye

Received: March 31, 2026

Accepted: May 12, 2026

Epub: May 18, 2026

ABSTRACT

Objectives: This study aimed to compare the discrimination and calibration performances of the Society of Thoracic Surgeons (STS) PROM and European System for Cardiac Operative Risk Evaluation II (EuroSCORE II) risk models in predicting early postoperative mortality and to evaluate the prognostic value of preoperative clinical factors in patients undergoing isolated coronary artery bypass grafting (CABG).

Methods: Sixty-four consecutive patients (mean age: 62.1 ± 10.1 years; 71.9% male) who underwent isolated CABG were included in this retrospective, single-center study. The discriminative capacity of the models was evaluated using receiver operating characteristic (ROC) curve analysis and the DeLong test, while calibration was assessed using observed/expected (O/E) mortality ratios.

Results: The overall operative mortality was 21.9% ($n=14$), demonstrating a statistically significant increase with surgical urgency (2.8% in the elective group; $p<0.001$). Both models underestimated actual mortality, particularly in emergency and salvage cases (O/E ratios: EuroSCORE II 24.7, STS PROM 21.9). In comparative group analyses, preoperative leukocyte (white blood cell) levels were significantly higher in the mortality group (11.95 ± 3.46 vs. 9.15 ± 2.74 , $p=0.006$). ROC analysis revealed similar discriminatory power for both models (area under the curve: STS PROM 0.749, EuroSCORE II 0.740; $p=0.866$). However, the sensitivity of the STS PROM model (71.4%) was higher than that of EuroSCORE II (57.1%).

Conclusion: In our small, single-center exploratory cohort undergoing isolated CABG, surgical urgency and elevated leukocyte levels were observed as potential clinical parameters associated with early mortality in comparative group analyses. Although the STS PROM and EuroSCORE II models demonstrated acceptable discriminatory capacity, they tend to underestimate operative mortality in cases of high surgical urgency. Nevertheless, for regional centers managing high-risk profiles, the STS PROM model may represent a more practical option when prioritizing sensitivity. These findings are strictly hypothesis-generating and warrant validation in larger, multicenter cohorts.

Keywords: Coronary artery bypass grafting, EuroSCORE, STS risk score, operative mortality, risk stratification.

Currently, various risk stratification models are utilized to predict operative mortality following open-heart surgery. Among these, the two most widely accepted and extensively validated models are the Society of Thoracic Surgeons (STS) score, which is based on North American databases, and the European System for Cardiac Operative Risk Evaluation II (EuroSCORE II), derived from European population data.^[1,2] Both models employ complex, multivariable algorithms to calculate expected mortality rates, thereby providing surgeons with objective prognostic data.

However, the predictive performance of these risk scoring systems can vary significantly depending on the geographical region of application, population demographics, and the dynamics of the local healthcare system.^[3] Studies have reported discrepancies in the “calibration” (the agreement between expected and observed mortality) and “discrimination” (the ability to distinguish between patients who survive and those who die) of these models. While some studies suggest that EuroSCORE II is a superior predictor of mortality, others emphasize the dominance of the STS score; conversely, several reports indicate that both scores may either overestimate or underestimate the actual operative risk.^[4,5]

Address for Correspondence: Anıl Akbaş, MD, Clinic of Cardiovascular Surgery, Batman Training and Research Hospital, Batman, Türkiye

E-mail: dr.anilakbas@gmail.com **ORCID ID:** orcid.org/0000-0003-0507-3498

Cite this article as: Akbaş A, Cebe T, Balkaya Aİ. Limitations of risk stratification in CABG: calibration across elective and emergency cohorts at a regional center. *Cardiovasc Surg Int*. [Epub Ahead of Print]



In developing nations and among patient cohorts with distinct risk profiles, the local validation of these global metrics is of paramount importance for the accurate interpretation of surgical outcomes. Therefore, the primary objective of this study was to compare the predictive performance of the EuroSCORE II and STS PROM models in patients undergoing isolated coronary artery bypass grafting (CABG) at our center and to investigate the potential calibration defects these algorithms exhibit in clinical scenarios characterized by high surgical urgency.

METHODS

Study Design and Patient Selection

This retrospective, single-center, observational study was conducted at the Clinic of Cardiovascular Surgery, Batman Training and Research Hospital. The medical records of patients who underwent surgery at our clinic between January 2023 and January 2026 were retrospectively reviewed. Following the application of the exclusion criteria defined in the study protocol, 64 patients who underwent isolated CABG were included in the final study cohort. The study protocol was approved by the Clinical Research Local Ethics Committee of Batman Training and Research Hospital (decision no: 460, date: January 28, 2026), and the research was conducted in strict adherence to the principles of the Declaration of Helsinki. Written informed consent was obtained from all participants included in the study prior to their surgical procedures, including consent for the use of their anonymized medical data for research purposes.

Patients aged 18 years or older who underwent isolated CABG under elective or emergency conditions were included in the study. Patients requiring concomitant cardiac procedures (e.g., valve repair/replacement, aortic surgery) and those undergoing re-operations were excluded. Furthermore, patients receiving preoperative mechanical circulatory support (IABP/ECMO) were excluded from the cohort. This methodological decision was implemented to prevent artificial deviations in the expected mortality calculations and to preserve the standardization of the statistical comparison between the STS PROM and EuroSCORE II models, owing to fundamental structural discrepancies in how these two models categorize and score mechanical support.

A total of 69 consecutive patients scheduled for CABG surgery at our clinic during the study period were initially evaluated. According to the predefined exclusion criteria in the study protocol, 5 patients (7.2%) who required concomitant cardiac procedures or received preoperative mechanical circulatory support systems were excluded. Following the application of these exclusion criteria, 64 patients who underwent isolated CABG were included in the final analysis.

Data Collection and Risk Scoring

The patients' demographic characteristics, preoperative clinical status, laboratory parameters, operative data, and postoperative outcomes were retrospectively extracted from the hospital data management system and patient medical records. Using preoperative data, the expected operative mortality risk percentages for each patient were calculated using the online calculators for EuroSCORE II (<http://www.euroscore.org>) and STS PROM (<http://riskcalc.sts.org>). The primary endpoint of the study was defined as operative mortality. Operative mortality was defined as in-hospital death or any-cause mortality occurring within the first 30 days post-discharge.

Surgical Strategy and Myocardial Protection

All procedures were performed via median sternotomy, utilizing standard cardiopulmonary bypass (CPB). As the revascularization strategy, the left internal mammary artery was routinely used to revascularize the left anterior descending artery, while saphenous vein grafts were preferred for other target vessels. Myocardial protection was achieved with isothermal blood cardioplegia across the entire cohort. In all cases involving saphenous vein grafts, proximal aortic anastomoses were routinely performed using a side-clamp technique.

Definition of Surgical Urgency Status and Calibration Analysis

The surgical urgency status of the included patients was classified into four categories in accordance with standard definitions from international cardiovascular surgery databases (STS and EuroSCORE): (1) Elective: Patients operated on a routine, scheduled basis without medical necessity for early intervention. (2) Urgent: Patients admitted via the emergency department or outpatient clinic who, for clinical reasons (e.g., unstable angina pectoris or critical coronary anatomy), underwent surgery during the same admission period without being discharged. (3) Emergency: Patients whose clinical status (e.g., hemodynamic instability or refractory ischemia) mandated surgical intervention before the beginning of the next routine working day. (4) Salvage: Patients in a life-threatening condition, such as cardiogenic shock, requiring emergency surgical intervention while intubated or under ongoing cardiopulmonary resuscitation.

The calibration performance of the risk scoring models was analyzed by calculating the observed-to-expected (O/E) mortality ratios across the entire cohort and within the aforementioned surgical urgency subgroups. An O/E ratio of 1.0 indicates perfect calibration, whereas values greater than 1.0 indicate that the model underestimates mortality.

Statistical Analysis

Statistical analyses of the study data were performed using R software (R Core Team, Vienna, Austria). The conformity of continuous variables to a normal distribution was evaluated using the Shapiro-Wilk test. The Mann-Whitney U test was used to compare continuous variables between two independent groups, while Fisher's exact test was preferred for analyzing the relationships between categorical variables. In the evaluation of operative times, because the limited number of cases in the subgroups would increase the risk of Type II error, no independent p-values were calculated for comparisons between these subgroups, and the data were presented solely as descriptive statistics. Analytical comparisons for these parameters were performed under non-parametric assumptions, using the Mann-Whitney U test to compare exclusively the "Elective" and "Emergency/Complicated" main groups.

The discriminative ability of the EuroSCORE II and STS PROM predictive models for forecasting mortality was evaluated using receiver operating characteristic (ROC) curve analysis, and the area under the curve (AUC) with its 95% confidence interval (CI) was calculated. The DeLong test was used to compare the AUCs of the two independent risk models. For model calibration, it was anticipated that the Hosmer-Lemeshow goodness-of-fit test would carry a high risk of violating asymptotic assumptions because of the limited sample size of the study cohort (n=64) and the low number of expected events (<5) in specific risk deciles. Therefore, this test was not applied. Instead, calibration performance was evaluated directly through O/E mortality ratios and

their corresponding 95% CIs. In all analyses, a two-tailed p-value of less than 0.05 was considered statistically significant.

RESULTS

Baseline Demographic and Clinical Characteristics

The mean age of the 64 patients included in the study was 62.1 ± 10.1 years (range: 41-81), comprising 71.9% (n=46) males and 28.1% (n=18) females. Analysis of the preoperative comorbidity profile revealed the following comorbidities: Hypertension in 45.3% (n=29), diabetes mellitus (DM) in 26.6% (n=17), chronic obstructive pulmonary disease in 7.8% (n=5), previous cerebrovascular accident (CVA) in 6.2% (n=4), and chronic kidney disease (CKD) in 3.1% (n=2). The mean preoperative ejection fraction of the entire population was $53.5 \pm 8.7\%$, and the mean serum creatinine level was 0.94 ± 0.24 mg/dL.

Analysis of Mortality and Associated Factors

Early postoperative mortality was observed in 21.9% (n=14) of the patients, while 78.1% (n=50) survived to discharge. When comparing the mortality (n=14) and survivor (n=50) groups, no statistically significant differences were detected with respect to age, gender, or the presence of DM, HT, CKD, or CVA ($p > 0.05$). Quantitative analysis of preoperative hematological and biochemical parameters showed that although serum creatinine levels were relatively higher in the mortality group (1.03 ± 0.29 mg/dL) compared to the survivor group (0.91 ± 0.22 mg/dL), this difference did not reach statistical significance ($p = 0.154$). Similarly, hematocrit levels (mortality: $43.13 \pm 4.74\%$ vs. survivor: $42.74 \pm 4.95\%$, $p = 0.789$) and platelet counts (mortality: $240.00 \pm 40.96 \times 10^3/\mu\text{L}$ vs. survivor: $244.76 \pm 79.38 \times 10^3/\mu\text{L}$, $p = 0.763$) demonstrated a homogeneous distribution between the two groups. In contrast, the preoperative white blood cell (WBC) count, a primary biomarker of the systemic inflammatory response, was significantly elevated in the mortality cohort ($11.95 \pm 3.46 \times 10^3/\mu\text{L}$) compared to the surviving cohort ($9.15 \pm 2.74 \times 10^3/\mu\text{L}$) ($p = 0.006$) (Table 1).

Analysis of CPB and Cross-clamp Times Across Surgical Urgency Subgroups

Evaluation of operative durations across the entire study cohort (n=64) according to surgical urgency revealed that patients undergoing elective surgery (n=36) had a mean CPB time of 92.5 minutes and a mean aortic cross-clamp time of 50.8 minutes (Table 2). In patients operated on an urgent basis (n=19), the mean CPB time was 118.3 minutes (maximum: 208 min), and the mean cross-clamp time was 72.5 minutes (maximum: 133 min). In the emergency group (n=7), the mean CPB and cross-clamp times were 99.1 and 50.3 minutes, respectively. For cases in the salvage category (n=2), the mean CPB time was 139.0 minutes, and the mean cross-clamp time was 69.0 minutes.

When the cohort was dichotomized into two main categories: Elective (n=36) and emergency/complicated (urgent, emergency, and salvage; n=28) cases, the mean cross-clamp time in the emergency/complicated group (66.7 min) was significantly longer than in the elective cases (50.8 min) ($p = 0.037$).

Calibration Performance of Risk Models Stratified by Surgical Urgency

Across the overall population, the mean expected mortality rates were calculated as $0.88 \pm 0.58\%$ for EuroSCORE II and $1.00 \pm 0.84\%$ for STS PROM. The actual observed mortality was 21.9%, yielding overall cohort O/E ratios of 24.7 for EuroSCORE II and 21.9 for STS PROM. The calibration analysis findings, stratified by surgical urgency status, are as follows (Table 3).

- **Elective cases (n=36):** While the observed mortality was 2.8% (n=1), the expected mortality was calculated as 0.71% for EuroSCORE II and 0.74% for STS PROM; the O/E ratios were determined to be 3.9 for EuroSCORE II and 3.7 for STS PROM.
- **Urgent cases (n=19):** With an observed mortality of 31.6% (n=6), the expected mortality rates were calculated at 0.82% for EuroSCORE II and 0.96% for STS PROM, resulting in O/E ratios of 38.3 and 33.0, respectively.

Table 1. Distribution of demographic and clinical variables according to early mortality

Variables	Total (n=64)	Survivors (n=50)	Mortality (n=14)	p-value
Age (years)	62.12±10.05	61.30±9.32	65.07±12.24	0.300
Male sex, n (%)	46 (71.9%)	37 (74.0%)	9 (64.3%)	0.512
Diabetes mellitus, n (%)	17 (26.6%)	11 (22.0%)	6 (42.9%)	0.170
Hypertension, n (%)	29 (45.3%)	24 (48.0%)	5 (35.7%)	0.608
CKD, n (%)	2 (3.1%)	1 (2.0%)	1 (7.1%)	0.392
COPD, n (%)	5 (7.8%)	2 (4.0%)	3 (21.4%)	0.065
Previous CVA, n (%)	4 (6.2%)	3 (6.0%)	1 (7.1%)	1.000
EF (%)	53.52±8.67	53.90±8.22	52.14±10.32	0.565
WBC ($10^3/\mu\text{L}$)	9.76±3.11	9.15±2.74	11.95±3.46	0.006*
HCT (%)	42.82±4.87	42.74±4.95	43.13±4.74	0.789
PLT ($\times 10^3/\mu\text{L}$)	243.72±72.47	244.76±79.38	240.00±40.96	0.763
Creatinine (mg/dL)	0.94±0.24	0.91±0.22	1.03±0.29	0.154
STS score	1.00±0.84	0.81±0.46	1.66±1.42	0.005*
EuroSCORE II	0.88±0.58	0.75±0.28	1.36±1.01	0.006*

Continuous variables are presented as mean \pm standard deviation; categorical variables are presented as absolute frequencies (n) and column percentages (%). EF: Ejection fraction; CKD: Chronic kidney disease; COPD: Chronic obstructive pulmonary disease; CVA: Cerebrovascular accident; WBC: White blood cell; HCT: Hematocrit; PLT: Platelet; STS: Society of Thoracic Surgeons; *: Indicates statistical significance ($p < 0.05$).

Table 2. Statistical comparison of cardiopulmonary bypass and aortic cross-clamp times according to surgical urgency status

Surgical urgency	Patients (n)	CPB time (min)	p-value	Cross-clamp time (min)	p-value
Elective cases	36	92.5 (56-165)	0.061	50.8 (23-99)	0.037*
Emergency and complicated cases	28	115.0 (56-208)		66.7 (31-133)	
-Urgent	19	118.3 (56-208)	-	72.5 (31-133)	-
-Emergency	7	99.1 (67-144)	-	50.3 (33-75)	-
-Salvage	2	139.0 (80-198)	-	69.0 (51-87)	-

CPB: Cardiopulmonary bypass; min: Minutes. Data are presented as mean (minimum-maximum). Statistical analysis was performed exclusively between the two main categories ("Elective" and "Emergency and Complicated") using the Mann-Whitney U test; *: Indicates statistical significance (p<0.05).

Table 3. Calibration analysis of risk models according to surgical urgency status

Surgical urgency	Number of patients, n (%)	Observed mortality (%)	Expected EuroSCORE II (%)	Expected STS PROM (%)	EuroSCORE II O/E ratio (95% CI)*	STS PROM O/E ratio (95% CI)
Elective	36 (56.3%)	2.8% (n=1)	0.71%	0.74%	3.9 (0.1-21.8)	3.7 (0.1-20.9)
Urgent	19 (29.7%)	31.6% (n=6)	0.82%	0.96%	38.3 (14.1-83.8)	33.0 (12.1-71.6)
Emergency	7 (10.9%)	71.4% (n=5)	1.79%	1.57%	39.9 (13.0-93.1)	45.4 (14.8-106.2)
Salvage	2 (3.1%)	100.0% (n=2)	1.36%	3.97%	73.3 (8.9-265.6)	25.2 (3.1-91.0)
Overall cohort	64 (100%)	21.9% (n=14)	0.88%	1.00%	24.7 (13.6-41.7)	21.9 (12.0-36.7)

O/E: Observed/Expected; CI: Confidence interval, STS: Society of Thoracic Surgeons; EuroSCORE II: European System for Cardiac Operative Risk Evaluation II.

- **Emergency cases (n=7):** The observed mortality was 71.4% (n=5), whereas the expected mortality was calculated as 1.79% for EuroSCORE II and 1.57% for STS PROM, yielding O/E ratios of 39.9 and 45.4, respectively.
- **Salvage cases (n=2):** Although a 100% observed mortality (n=2) was recorded, the expected risk scores were computed as 1.36% for EuroSCORE II and 3.97% for STS PROM; the resulting O/E ratios were 73.3 for EuroSCORE II and 25.2 for STS PROM.

Analysis of risk scores evaluating the preoperative risk stratification scores, the predicted risk scores of the patients who died were found to be significantly higher than those of the survivor group.

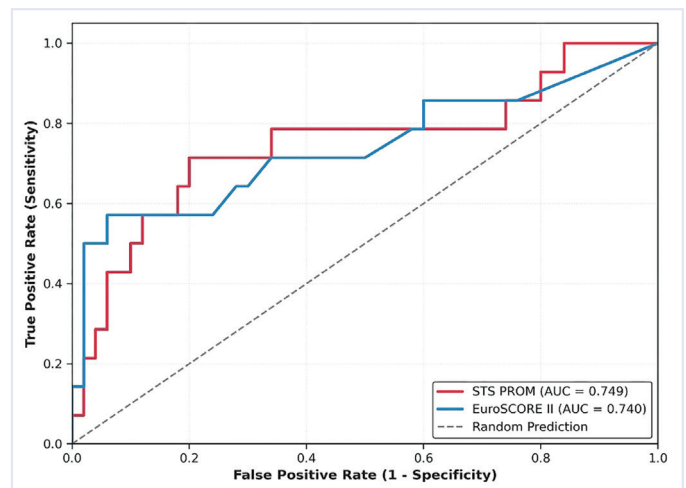
- **EuroSCORE II:** Calculated as a mean of 1.36 ± 1.01 in the mortality group compared to 0.75 ± 0.28 in the survivor group ($p=0.006$).
- **STS PROM:** Calculated as a mean of $1.66 \pm 1.42\%$ in the mortality group versus $0.81 \pm 0.46\%$ in the survivor group ($p=0.005$).

ROC analysis and predictive performance according to the ROC curve analysis conducted to evaluate the discriminative performances of the risk models in predicting mortality:

- The AUC for STS PROM was calculated as 0.749 (95% CI: 0.569-0.905),
- The AUC for EuroSCORE II was calculated as 0.740 (95% CI: 0.547-0.910) (Figure 1).

When the AUC values of the two risk models were compared using the DeLong test, no statistically significant difference was detected between them ($p=0.866$). The optimal cut-off points for both models were determined based on the maximum Youden index ($J = \text{Sensitivity} +$

Specificity - 1) from the curve coordinates. Based on these calculations, the maximum Youden index ($J=0.514$) for the STS PROM model was identified at a threshold score >1.04 . At this threshold, sensitivity and specificity were 71.4% and 80.0%, respectively ($p<0.05$). For the EuroSCORE II model, the maximum Youden index ($J=0.511$) was established at a score >1.17 , yielding a sensitivity of 57.1% and a specificity of 94.0% ($p<0.05$).

**Figure 1.** ROC curves demonstrating the discriminative performances of the STS PROM and EuroSCORE II risk models in predicting early postoperative mortality.

ROC: Receiver operating characteristic; STS: Society of Thoracic Surgeons; AUC: Area under the curve; EuroSCORE II: European System for Cardiac Operative Risk Evaluation II.

DISCUSSION

In this study investigating the performance of the STS PROM and EuroSCORE II risk models in patients undergoing isolated CABG, surgical urgency had a marked effect on postoperative mortality. Our findings indicate that both models have a statistically acceptable discriminative ability for predicting mortality. However, no statistically significant differences were detected in the overall performance of the models. When evaluating model performance in our study cohort, the STS PROM model demonstrated greater sensitivity for identifying high-risk patients, whereas the EuroSCORE II model appeared more selective for identifying low-risk patients. This indicates that neither model provides absolute superiority over the other; rather, they offer different advantages in clinical application.

A review of the cardiovascular surgery literature reveals that mortality rates secondary to coronary artery bypass surgery range within a narrow band of 1-3% in stable elective cases,^[6-8] whereas this rate can rise to 20-50% in patients operated on under emergency and salvage status due to cardiogenic shock or ongoing myocardial ischemia.^[9] The overall operative mortality rate of 21.9% observed in our study is notably higher than that reported in standard elective series in the literature. This elevated operative mortality stems from the referral to our center of patients who could not be operated on at surrounding healthcare facilities because of high surgical risk or technical limitations. Consequently, this situation significantly increases the proportion of patients in our cohort who are taken directly to emergency surgery without the opportunity for preoperative stabilization. Therefore, the presented data do not reflect a standard, stable series of elective patients; rather, they represent the clinical outcomes of a specific high-risk patient population referred to our clinic because of regional circumstances.

The 2.8% mortality rate observed in elective cases in our study was consistent with global databases reported in the literature. In contrast, high mortality rates were recorded: 31.6% in the urgent group, 71.4% in the emergency group, and 100% in the salvage group. The elevated mortality observed in the emergency and salvage categories is associated with the patients' acute hemodynamic decompensation and limited physiological reserves during the perioperative period. This group comprises patients undergoing obligatory surgery against a background of cardiogenic shock or refractory ischemia, without the opportunity for preoperative medical optimization. The advanced comorbid burden and acute cardiovascular collapse diminish the capacity to compensate for additional systemic stress induced by surgical trauma and CPB, thereby increasing early mortality.

Although the mortality observed in the two cases operated on for salvage indications is not suitable for statistical generalization due to the limited sample size, it provides a clinical basis for explaining the prediction deviations of the models in high-risk groups. In patients undergoing surgery while in cardiogenic shock or receiving active resuscitation, the static parameters in standard risk-scoring systems may fail to fully reflect acute hemodynamic collapse. The high mortality trend observed in emergency and salvage operations may be related to dynamic pathophysiological processes, such as acute myocardial ischemia, profound systemic hypoperfusion, and cellular acidosis, which cannot be adequately incorporated into the mathematical scoring of the models in question.

In the performance evaluation of the risk scoring systems, the STS PROM and EuroSCORE II demonstrated statistically similar overall discriminative performance. The finding that the AUC values for both models were above the 0.70 threshold indicates that these scoring systems possess an acceptable level of discrimination in the examined cohort. Indeed, the analysis of the ROC curves using the DeLong test revealed no statistically significant difference between the two models. Model calibration was assessed descriptively using O/E ratios, and subgroup calibration estimates remain inherently unstable, with high variance due to very small sample sizes within these urgency strata.

When evaluating the clinical utility of risk models, sensitivity and specificity at pre-specified cut-off points are important alongside overall AUC values. Although the overall discriminatory capacities of both models were similar in our study, the sensitivity of the STS PROM model for predicting mortality (71.4%) was higher than that of EuroSCORE II (57.1%). This finding indicates that the STS PROM model may be more sensitive in detecting high-risk patients within the examined cohort, thereby reducing the probability of false negatives. On the other hand, the EuroSCORE II model, which exhibited a specificity of 94.0%, was more selective in distinguishing low-risk patients.

These performance differences between the two models may be related to the geographic origins and structural characteristics of the databases from which the algorithms were derived. Although both scoring systems were derived from large cardiac surgery populations, including valve and combined procedures, EuroSCORE II is based on a European-centric patient profile, whereas STS PROM originates from a North American database and incorporates much more detailed clinical variables in its calculation tool. Indeed, it has been reported in the literature that the EuroSCORE II model tends to underestimate mortality, particularly in the highest-risk patient quartile.^[10] In our cohort, which predominantly comprises patients with severe clinical presentations referred to our clinic because of regional circumstances, the STS PROM model, which uses a more detailed set of variable parameters, was observed to predict mortality more accurately and to exhibit superior predictive performance in the high-risk group.

Another from our comparative group analyses is that preoperative leukocyte (WBC) levels were significantly higher in the group in which operative mortality was observed. The observed leukocytosis may indicate a systemic stress response secondary to acute myocardial injury or underlying subclinical inflammation. Indeed, there are studies in the literature reporting that preoperative systemic inflammatory burden may adversely affect postoperative clinical outcomes in patients undergoing CABG.^[11] However, due to the retrospective nature of our study, the inability to evaluate more specific inflammatory markers such as C-reactive protein, procalcitonin, or neutrophil-to-lymphocyte ratio across the entire cohort, and the inability to completely rule out potential concomitant infectious pathologies constitute important limitations. Considering that multivariate analysis could not be performed due to the limited number of events, the detected preoperative WBC elevation should be evaluated as an observed association and as a "hypothesis-generating" finding for future, more comprehensive studies rather than as a definitive independent risk factor.

In isolated CABG, surgical urgency status, which is a primary indicator of preoperative physiological status, emerges as one of the important clinical determinants of early postoperative mortality. Our research findings indicate that although the STS PROM and EuroSCORE II risk

models may possess a general discriminative ability in this patient group, they tend to underestimate the actual mortality rate, particularly in emergency and salvage cases. In this context, for hemodynamically unstable high-risk cases, it is clinically more appropriate to approach with caution the predictive deviations that standard scoring systems may present, and to center clinical decision-making on the patient's immediate physiological decompensation rather than on algorithmic scores. Furthermore, findings such as preoperative leukocyte elevation, which were found to be associated with mortality in comparative group analyses, may serve as hypothesis-generating features to be evaluated in future large-scale, multivariate studies.

Certain methodological limitations should be considered when evaluating the results of our study. First, the single-center, retrospective design of the study may limit the direct applicability of the findings to the general population. Although the consecutive inclusion of patients contributed to reducing the risk of selection bias, it cannot entirely eliminate this risk.

Second, the most fundamental limitation from a biostatistical perspective is the limited total number of operative mortality events. In the context of statistical validity, constructing a reliable multivariate logistic regression model requires a minimum of 10 events per independent variable, in accordance with the "events per variable" principle. Because the number of events in our sample did not fully meet this statistical assumption, we avoided multivariate analysis to prevent model overfitting. This methodological necessity restricted the ability to clearly evaluate whether factors such as surgical urgency, advanced age, or elevated leukocyte levels are independent prognostic predictors of postoperative mortality.

Third, due to the restricted sample size across different urgency strata, detailed and statistically powered subgroup analyses could not be performed. The highly limited number of patients undergoing salvage surgery necessitates that the high mortality observed in this group be interpreted as a clinical observational trend rather than definitive statistical evidence, which restricts the deep interpretability of subgroup-specific outcomes. Consequently, the data obtained are exploratory in nature; it is considered that multicenter and prospective studies with higher statistical power may be required to validate the independent prognostic value of these observed associations and the calibration performances of the models in high-risk groups.

Ethics

Ethics Committee Approval: The study protocol was approved by the Clinical Research Local Ethics Committee of Batman Training and Research Hospital (decision no: 460, date: January 28, 2026), and the research was conducted in strict adherence to the principles of the Declaration of Helsinki.

Informed Consent: Written informed consent was obtained from all participants included in the study prior to their surgical procedures, including consent for the use of their anonymized medical data for research purposes.

Footnotes

Authorship Contributions

Surgical and Medical Practices: A.A., T.C., A.İ.B.; Concept: A.A.; Design: A.A., T.C.; Data Collection or Processing: A.A., A.İ.B.; Analysis or Interpretation: A.A.; Literature Search: A.A., T.C., A.İ.B.; Writing: A.A.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study received no financial support.

REFERENCES

- Jacobs JP, Shahian DM, Badhwar V, Thibault DP, Thourani VH, Rankin JS, et al. The Society of Thoracic Surgeons 2021 adult cardiac surgery risk models for multiple valve operations. *Ann Thorac Surg.* 2022;113:511-8. doi: 10.1016/j.athoracsur.2021.03.089.
- Nashef SA, Roques F, Sharples LD, Nilsson J, Smith C, Goldstone AR, et al. EuroSCORE II. *Eur J Cardiothorac Surg.* 2012;41:734-44; discussion 744-5. doi: 10.1093/ejcts/ezs043.
- Wolf PJW, Amato VL. Assessment of EuroSCORE II and STS score performance and the impact of surgical urgency in isolated coronary artery bypass graft surgery at a referral center in São Paulo, Brazil. *Braz J Cardiovasc Surg.* 2024;39:e20230282. doi: 10.21470/1678-9741-2023-0282.
- Kunt AG, Kurtcephe M, Hidiroglu M, Cetin L, Kucuker A, Bakuy V, et al. Comparison of original EuroSCORE, EuroSCORE II and STS risk models in a Turkish cardiac surgical cohort. *Interact Cardiovasc Thorac Surg.* 2013;16:625-9. doi: 10.1093/icvts/ivt022.
- Shapira-Daniels A, Blumenfeld O, Korach A, Rudis E, Izhar U, Shapira OM. The American Society of Thoracic Surgery Score versus EuroSCORE I and EuroSCORE II in Israeli patients undergoing cardiac surgery. *Isr Med Assoc J.* 2019;21:671-5.
- Chua TKT, Gao F, Chia SY, Sin KYK, Naik MJ, Tan TE, et al. Long-term mortality after isolated coronary artery bypass grafting and risk factors for mortality. *J Cardiothorac Surg.* 2024;19:429. doi: 10.1186/s13019-024-02943-0.
- Ferguson TB. Mortality in coronary artery bypass grafting: what's next? *Circulation.* 2012;125:2409-11. doi: 10.1161/CIRCULATIONAHA.112.106856.
- Shahian DM, O'Brien SM, Sheng S, Grover FL, Mayer JE, Jacobs JP, et al. Predictors of long-term survival after coronary artery bypass grafting surgery: results from the Society of Thoracic Surgeons Adult Cardiac Surgery Database (the ASCERT study). *Circulation.* 2012;125:1491-500. doi: 10.1161/CIRCULATIONAHA.111.066902.
- Santarpino G, Ruggieri VG, Mariscalco G, Bounader K, Beghi C, Fischlein T, et al. Outcome in patients having salvage coronary artery bypass grafting. *Am J Cardiol.* 2015;116:1193-8. doi: 10.1016/j.amjcard.2015.07.034.
- Patrat-Delon S, Rouxel A, Gacouin A, Revest M, Flécher E, Fouquet O, et al. EuroSCORE II underestimates mortality after cardiac surgery for infective endocarditis. *Eur J Cardiothorac Surg.* 2016;49:944-51. doi: 10.1093/ejcts/ezv223.
- Walker AE, Seibert SM, Donato AJ, Pierce GL, Seals DR. Vascular endothelial function is related to white blood cell count and myeloperoxidase among healthy middle-aged and older adults. *Hypertension.* 2010;55:363-9. doi: 10.1161/HYPERTENSIONAHA.109.145870.