Comparison of Two Major Perioperative Bleeding Scores for Cardiac Surgery Trials

Universal Definition of Perioperative Bleeding in Cardiac Surgery and European Coronary Artery Bypass Grafting Bleeding Severity Grade

Justyna Bartoszko, M.D., Duminda N. Wijeysundera, M.D., Keyvan Karkouti, M.D., on behalf of the Transfusion Avoidance in Cardiac Surgery Study Investigators*

ABSTRACT

Background: Research into major bleeding during cardiac surgery is challenging due to variability in how it is scored. Two consensus-based clinical scores for major bleeding: the Universal definition of perioperative bleeding and the European Coronary Artery Bypass Graft (E-CABG) bleeding severity grade, were compared in this substudy of the Transfusion Avoidance in Cardiac Surgery (TACS) trial.

Methods: As part of TACS, 7,402 patients underwent cardiac surgery at 12 hospitals from 2014 to 2015. We examined content validity by comparing scored items, construct validity by examining associations with redo and complex procedures, and criterion validity by examining 28-day in-hospital mortality risk across bleeding severity categories. Hierarchical logistic regression models were constructed that incorporated important predictors and categories of bleeding.

Results: E-CABG and Universal scores were correlated (Spearman ρ = 0.78, $P < 0.0001$), but E-CABG classified 910 (12.4%) patients as having more severe bleeding, whereas the Universal score classified 1,729 (23.8%) as more severe. Higher E-CABG and Universal scores were observed in redo and complex procedures. Increasing E-CABG and Universal scores were associated with increased mortality in unadjusted and adjusted analyses. Regression model discrimination based on predictors of perioperative mortality increased with additional inclusion of the Universal score (c-statistic increase from 0.83 to 0.91) or E-CABG (c-statistic increase from 0.83 to 0.92). When other major postoperative complications were added to these models, the association between Universal or E-CABG bleeding with mortality remained.

Conclusions: Although each offers different advantages, both the Universal score and E-CABG performed well in the validity assessments, supporting their use as outcome measures in clinical trials. **(ANESTHESIOLOGY 2018; 129:1092-100)**

EXCESSIVE perioperative blood loss requiring transfusion is a common and clinically important complication of cardiac surgery.[1,2] Despite its importance, there is significant variability in how this outcome is scored across clinical trials.[1,2] Adoption of consensus-based outcome scoring methods has been advocated as a means for better standardizing endpoints in clinical trials.[3–6] They offer consistency across clinical trials, help simplify the interpretation of trials with conflicting results, and facilitate evidence synthesis.[7,8]

Two such consensus-based scores for clinically important blood loss in cardiac surgery have recently been proposed, namely the Universal definition of perioperative bleeding[9]

Editor's Perspective

What We Already Know about This Topic

- Major bleeding can occur during cardiac surgery. Although different scoring systems exist, the assessment of bleeding can be variable, and the reliability of these scoring systems has not been determined.

What This Article Tells Us That Is New

- Two consensus-based scoring systems for assessing bleeding were compared in a substudy of the Transfusion Avoidance in Cardiac Surgery trial. Both the Universal score and European Coronary Artery Bypass Graft scores performed well and may be used as validated outcome measures in future clinical trials.

Table 1. Bleeding Categories Defined by Two Different Consensus-based Bleeding Scores Designed for Use in Cardiac Surgery Clinical Trials

	Sternal Closure Delayed	Postoperative Chest Tube Blood Loss within 12 h (ml)	Erythrocytes (units)	FFP (units)	PLT (units)	Cryoprecipitate	PCCs	rFVIIa	Reexploration/ Tamponade
Universal Definition of Perioperative Bleeding in Adult Cardiac Surgery*[9]									
Bleeding class									
Class 0 (insignificant)	No	< 600	0*	0	0	No	No	No	No
Class 1 (mild)	No	601–800	1	0	0	No	No	No	No
Class 2 (moderate)	No	801–1000	2–4	2–4	Yes	Yes	Yes	No	No
Class 3 (severe)	Yes	1001–2000	5–10	5–10	N/A	N/A	N/A	No	Yes
Class 4 (massive)	N/A	> 2000	> 10	> 10	N/A	N/A	N/A	Yes	N/A
E-CABG grading and additive score†[10]									
Grades		Intervention for Treatment of Bleeding							
Grade 0		No use of blood products with the exception of 1 unit of erythrocytes							
Grade 1		Transfusion of platelets; transfusion of fresh frozen plasma or prothrombin complex concentrates; Transfusion of 2–4 units of erythrocytes							
Grade 2		Transfusion of 5–10 units of erythrocytes; reoperation for bleeding							
Grade 3		Transfusion of >10 units of erythrocytes							

*If clinical events occurring to a patient indicate more than one bleeding category, the worst category applies, including if transition between categories is based on one more severe item. The Universal definition of perioperative bleeding is scored based on events occurring during surgery or within the first postoperative day. Preoperative transfusions are not included. †This classification included any transfusion of erythrocytes, platelets, and fresh frozen plasma that occurred during the operation and postoperatively during the same in-hospital stay. Preoperative transfusions are not included.

E-CABG, European Coronary Artery Bypass Graft; FFP, fresh frozen plasma; N/A, not applicable; PCC, prothrombin complex concentrate; PLT, platelet concentrate; rFVIIa, recombinant activated factor VII.

and the European Coronary Artery Bypass Graft (E-CABG) bleeding severity grade (table 1).[10,11] The Universal score is based on nine clinically important events that occur during surgery or within the first postoperative day. It was designed to capture significant bleeding independent of its source or clinical management decisions (table 1).[9] Conversely, the E-CABG score is based on interventions that indirectly quantify perioperative blood loss (*i.e.*, blood product transfusion, reoperation for bleeding) and occur at any point from surgery to the end of hospitalization (table 1).[11]

Precise quantification of bleeding during the perioperative period is difficult and prone to error. Direct estimation of surgical blood loss, even by experienced staff, is often inaccurate and unreliable.[12,13] In addition, ongoing bleeding may be concealed or unrecognized until it leads to changes in hemodynamic stability or laboratory parameters. As such, these consensus-based scores were designed to grade bleeding using multiple items rather than relying on directly observed blood loss. Recognizing that multiple items are combined to create a construct for bleeding, such scores are often compared and evaluated to provide evidence of their validity prior to use in clinical trials.[14–16]

There has been no prior comparison of E-CABG and the Universal score for their intended purpose, which is as clinical trial endpoints. We therefore conducted a substudy of the Transfusion Avoidance in Cardiac Surgery (TACS) trial,[17] which was a stepped-wedge clustered randomized controlled trial evaluating point-of-care hemostatic testing for transfusion avoidance at 12 hospitals, to compare the Universal and E-CABG scoring systems with respect to their content, construct, and criterion validity.

Content validity is whether all important domains of a given construct are included and was assessed by examining the individual items graded by the Universal score and E-CABG. Construct validity is whether a measurement tool captures the phenomenon it claims to measure and was assessed by examining whether patients who underwent procedures known to be associated with higher blood loss correspondingly had higher Universal and E-CABG scores. Last, criterion validity, which is the extent to which a score is related to an important outcome, was assessed by examining whether higher blood loss as assessed by either score was predictive of early postoperative mortality.

We hypothesize that both E-CABG and the Universal score will have evidence of content, criterion, and construct validity but expect that the Universal score may have better measurement properties overall given the greater number of clinically relevant events captured in the perioperative period.

Materials and Methods

Study Setting and Population

This substudy was based on data prospectively collected in the TACS trial, which included all patients who underwent elective, urgent, and emergent cardiac surgical procedures with cardiopulmonary bypass at 12 Canadian study sites from October 6, 2014, to May 1, 2015. Institutional research ethics board approval for this substudy was obtained from the University Health Network in Toronto, Canada. All authors had full access to all the data in the study and were responsible for its integrity and the data analysis. If a

patient was readmitted for additional operations requiring cardiopulmonary bypass during the study period, only data from the first admission were used. The eligible patient sample included all 7,402 patients included in the TACS trial.

Predictor and Outcome Definitions

The primary outcome was 28-day in-hospital all-cause mortality. In the TACS trial, major bleeding was scored using the Universal definition of perioperative bleeding, but E-CABG could be readily scored using variables already collected in the trial data set. Other events of interest included acute kidney injury (defined as an at least twofold postoperative increase in creatinine concentration or new need for renal replacement therapy, which corresponds to Kidney Disease Improving Global Outcomes stage 2 or 3),[18] sepsis, sternal infection, myocardial infarction, return to the operating room for reexploration, and cerebrovascular accident. In-hospital follow-up for complications was censored at postoperative day 28.

Important confounders for potential inclusion in risk-adjustment models were selected based on a review of published cardiac risk scores, as well as the literature examining the association between preoperative factors and postcardiac surgery outcomes.[19–29] These potential confounders included demographics (age, sex); procedure urgency (elective, urgent, emergent); redo procedure; preoperative intraaortic balloon pump; preoperative hemoglobin concentration (g/l); preoperative renal dysfunction, which was defined as an estimated glomerular filtration rate of less than 60 ml/min (calculated using the Cockcroft–Gault Equation)[30] or preoperative dialysis; diabetes mellitus; extracardiac arteriopathy, which was defined as stroke, transient ischemic events, or peripheral arterial disease; chronic obstructive pulmonary disease; coronary artery disease; liver disease; heart failure; hypertension; and procedure complexity.[20] Procedure complexity was classified as simple, which was defined as isolated coronary artery bypass graft (CABG) or single valve procedure, or complex (all other procedures).

Statistical Analysis

The data set had at most 2% missing data; hence, missing data were not replaced or imputed. Initially, descriptive statistics (mean, SD, median, interquartile range, counts, proportion) were used to characterize the overall cohort, strata based on the presence *versus* absence of the primary outcome, and strata delineated by the E-CABG and Universal scores. Statistical significance was defined by a two-sided *P* value less than 0.05. All statistical analyses were conducted using SAS version 9.4 (SAS Institute Inc., USA).

Content Validity. Content validity refers to whether a measure reasonably represents all aspects of a given construct. The items graded by each score, as well as items incorporated into one score but not the other, were assessed for how they contributed to assigning patients to a given category of blood loss. Chest tube output was specifically examined because it

could upgrade bleeding severity in the Universal score alone and thereby could result in discrepancies between the Universal score and E-CABG. The correlation between E-CABG and Universal scores was characterized using the Spearman statistic.

Construct Validity. Construct validity refers to whether the Universal score and E-CABG behave as would be expected if they are representing measures of perioperative blood loss. We examined whether procedure types known to have higher rates of blood loss, such as complex and redo procedures, were associated with higher blood loss severity as graded by either score.

Criterion Validity. An important component of criterion validity is whether increased severity of blood loss (as assessed by the Universal score or E-CABG) is associated with future events, such as increasing postoperative mortality. We measured the adjusted association between the Universal score and E-CABG with in-hospital 28-day all-cause mortality using multivariable logistic regression modeling. We accounted for clustering by using hierarchical models that incorporated site random effects and patient-level factors as fixed effects. Both the Universal score and E-CABG were each treated as ordinal variables, with progressively increasing scores treated as higher severity categories. Model assumptions were verified, including linearity of continuous variables. Age was modeled using a b-spline method to provide a robust and flexible way of modeling nonlinearity to the logit. Hemoglobin was treated as an untransformed continuous variable because it did not demonstrate significant nonlinearity. All other variables were binary or categorical.

A series of nested models were constructed for both the Universal score and E-CABG. Model 1 was used as the baseline model that only incorporated parsimonious predictive variables, which were initially identified from the literature and published predictive indices. Bootstrapping was used to select relevant variables for inclusion in the model from this initial list. Random sampling with replacement was used to generate a sample of 5,000 in 300 bootstrapped replicates, and covariates included in more than 50% of bootstrapped replicates were retained. Site was empirically retained in the model. Model 2 included all predictor variables in model 1 plus Universal score categories. Model 3 incorporated all predictor variables in model 2 plus major perioperative complications as predictors. This process was then repeated for E-CABG bleeding severity grades. Model 4 was composed of all predictor variables in model 1 plus E-CABG bleeding severity grades. Model 5 incorporated all predictor variables in model 4 plus major perioperative complications. Model calibration was examined using the Hosmer–Lemeshow statistic, whereas discrimination was characterized using the area under the curve for the receiver-operating-characteristic curve.

Using SAS version 9.4, the Glimmix procedure was used to create hierarchical models. The output data from the Glimmix procedure was subsequently used in the logistic

procedure to obtain the area under the curve and receiver-operating-characteristic curve for each model. Using Glimmix output, the logistic procedure receiver-operating-characteristic contrast command was used to compare differences in area under the curve values between the models. The output data sets from the Glimmix procedure were ranked according to deciles, and the Hosmer–Lemeshow statistic was calculated for each hierarchical model. In addition, internal model validation was conducted for all adjusted models *via* bootstrapping to obtain the model optimism, which was subsequently used to adjust the c-statistic and obtain a 95% CI using the Harrell Optimism SAS macro.

Results

The primary outcome, 28-day in-hospital mortality, occurred in 190 (2.6%) patients. The characteristics of patients stratified by mortality are presented in table 2. The Universal score could be fully calculated for 7,281 (98.4%) patients, of whom 168 (2.3%) died. E-CABG could be fully calculated for 7,347 (99.3%) of patients, of whom 190 (2.6%) died.

Content Validity

Universal score classes and E-CABG grades were moderately correlated with each other (Spearman ρ = 0.78, P < 0.0001). Only 910 (12.4%) patients were classified as having more severe bleeding (grade 2 or 3) when E-CABG was used, whereas 1,729 (23.8%) of patients were classified as having more severe bleeding (class 3 or 4) by the Universal score. Individual items in each scale were evaluated to explain this discrepancy. In total, 857 (11.9%) patients were classified by E-CABG as having lower severity bleeding (E-CABG grades 0 and 1) but classified as higher severity bleeding by the Universal score (Universal score class 3 and 4). Out of these individuals, 700 (81.7%) patients were classified as higher severity bleeding by the Universal score based solely on high chest tube output. Mortality risk among patients classified into higher severity bleeding categories due to chest tube output alone was low, with 0.9% (n = 6) of these patients experiencing death. In contrast, patients classified as higher severity by the Universal score due to factors other than or in addition to chest tube output had much higher mortality, with 12.6% (n = 130) experiencing death. No patients classified as lower severity bleeding by the Universal score (class 0 to 2) were scored as higher bleeding by E-CABG (grade 2 and 3).

Construct Validity

There were higher numbers of complex and redo procedures within each increasing category of bleeding severity for both scores (fig. 1).

Criterion Validity

When blood loss severity increased as graded by either score, there was a corresponding increase in mortality (fig. 2). In unadjusted logistic regression analyses, increasing Universal

Table 2. Patient Characteristics Stratified by In-hospital 28-day All-cause Mortality

Variable	Patients Alive (n = 7,212)	Patients with Primary Outcome (Death) (n = 190)
Age	67 (59, 74)	72 (63, 80)
Sex (female)	1,791 (24.83)	64 (33.68)
Procedure urgency		
Elective	4,543 (63.57)	77 (41.40)
Urgent	2,165 (30.29)	61 (32.80)
Emergent	439 (6.14)	48 (25.81)
Redo operation	417 (5.84)	31 (16.49)
Preoperative IABP	131 (1.82)	17 (8.95)
Procedure complexity	2,050 (28.42)	108 (56.84)
Preoperative hemoglobin concentration (g/l)	136 (123, 146)	121 (105, 137)
Comorbidities		
Diabetes	2,395 (33.21)	73 (38.42)
Extracardiac arteriopathy	1,216 (16.86)	62 (32.63)
COPD	880 (12.20)	31 (16.32)
Renal dysfunction	1,767 (24.98)	104 (58.10)
Coronary artery disease	5,266 (73.02)	123 (64.74)
Heart failure	1,445 (20.04)	78 (41.05)
Liver disease	158 (2.19)	13 (6.84)
Hypertension	5,554 (77.01)	144 (75.79)
Perioperative events		
Reexploration/return to operating room	333 (4.62)	66 (34.74)
AKI	43 (0.61)	8 (4.76)
CVA	112 (1.55)	34 (17.89)
Sternal infection	96 (1.33)	10 (5.26)
Acute MI	28 (0.39)	10 (5.26)
Postoperative sepsis	75 (1.04)	31 (16.32)
Universal definition of perioperative bleeding category		
Class 0	2,196 (30.87)	3 (1.79)
Class 1	1,200 (16.87)	9 (5.36)
Class 2	2,124 (29.86)	20 (11.90)
Class 3	1,325 (18.63)	76 (45.24)
Class 4	268 (3.77)	60 (35.71)
E-CABG grade		
Grade 0	4,090 (57.15)	17 (8.95)
Grade 1	2,295 (32.07)	35 (18.42)
Grade 2	715 (9.99)	115 (60.53)
Grade 3	57 (0.80)	23 (12.11)

The values shown are frequencies (%) or medians (Q1, Q3).
AKI, acute kidney injury; COPD, chronic obstructive pulmonary disease; CVA, cerebrovascular accident; E-CABG, European Coronary Artery Bypass Graft; IABP, intraaortic balloon pump; MI, myocardial infarction.

and E-CABG scores were associated with increasing mortality (table 3). When used as the only predictor for the outcome of mortality, the Universal score and E-CABG did not differ in their discrimination, with the Universal score having an area under the curve of 0.84 (95% CI 0.81 to 0.87), and E-CABG having an area under the curve of 0.85 (95% CI 0.82 to 0.88). These two areas under the curve did not differ significantly (P = 0.25).
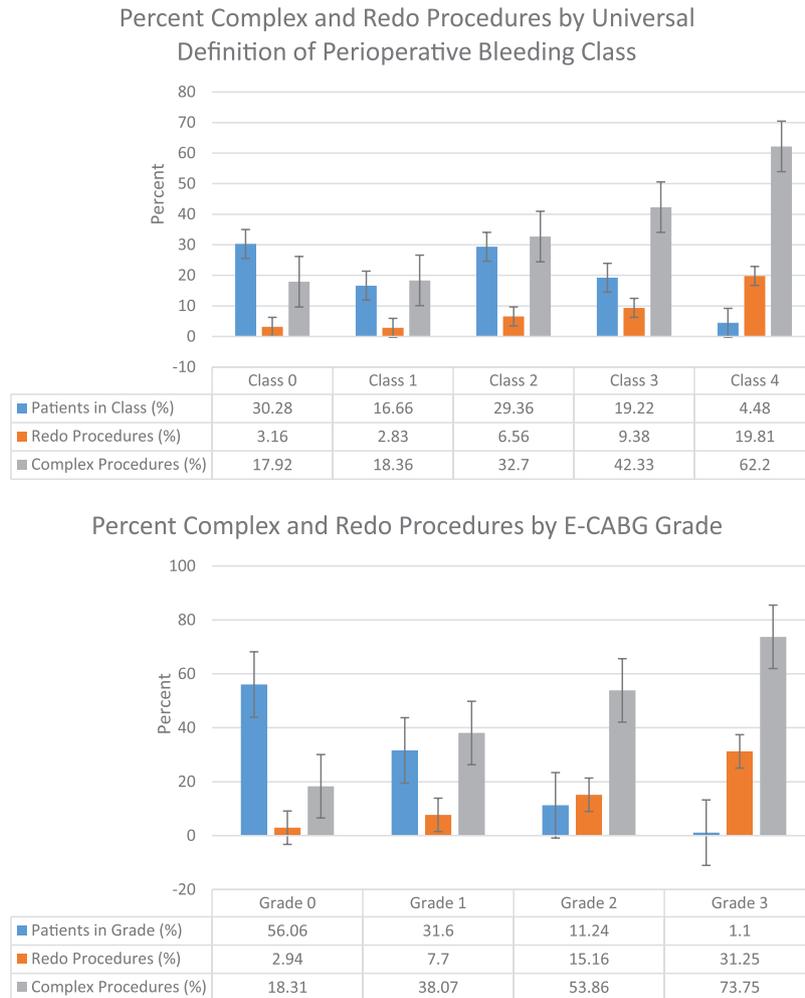
Percent Complex and Redo Procedures by Universal
Definition of Perioperative Bleeding Class

	Class 0	Class 1	Class 2	Class 3	Class 4
Patients in Class (%)	30.28	16.66	29.36	19.22	4.48
Redo Procedures (%)	3.16	2.83	6.56	9.38	19.81
Complex Procedures (%)	17.92	18.36	32.7	42.33	62.2

Percent Complex and Redo Procedures by E-CABG Grade

	Grade 0	Grade 1	Grade 2	Grade 3
Patients in Grade (%)	56.06	31.6	11.24	1.1
Redo Procedures (%)	2.94	7.7	15.16	31.25
Complex Procedures (%)	18.31	38.07	53.86	73.75

Fig. 1. Redo and complex procedure proportions shown for each bleeding severity category. E-CABG, European Coronary Artery Bypass Graft.

Similar results were found in subsequent adjusted analyses. In the hierarchical logistic regression model that incorporated important confounders but not bleeding events or other complications (model 1), the area under the curve for predicting 28-day in-hospital mortality was 0.84. In model 2, Universal classes were added, and the area under the curve increased to 0.91. In model 3, the further addition of other important complications increased the area under the curve to 0.94. Adding E-CABG grade as a predictor to model 1 similarly increased the area under the curve to 0.92. Further addition of other important complications to the model with E-CABG grades increased the area under the curve to 0.94. In each of the models, bleeding severity regardless of the score used, whether assessed by the Universal score or E-CABG, demonstrated a statistically significant association with 28-day in-hospital mortality. The details of the models can be found in Supplemental Digital Content 1, http://links.lww.com/ALN/B671. Detailed model calibration results can be found in Supplemental Digital Content 2, http://links.lww.com/ALN/B672.

Discussion

This study compared the Universal definition of perioperative bleeding and E-CABG bleeding severity scores and attempted to provide evidence of construct, criterion, and content validity in a high-quality data set from a multicenter clinical trial of patients undergoing cardiac surgery. Our findings indicate that the Universal score and E-CABG are both valid and acceptable scoring systems for use as bleeding endpoints in cardiac surgery clinical trials.

The Universal definition of perioperative bleeding has been studied in various patient samples since its development. It was tested in an 1,144 patient single-institution adult European cardiac surgical database, where it was shown that increasing classes demonstrated an independent association with 30-day mortality.[9] The Universal score was further validated in an institutional cardiac surgery data set of 2,764 patients in Finland, where increasing classes were significantly associated with worse immediate and late outcomes.[19]

E-CABG has also been studied in a variety of patient samples. It was evaluated in a 7,491 patient sample drawn from
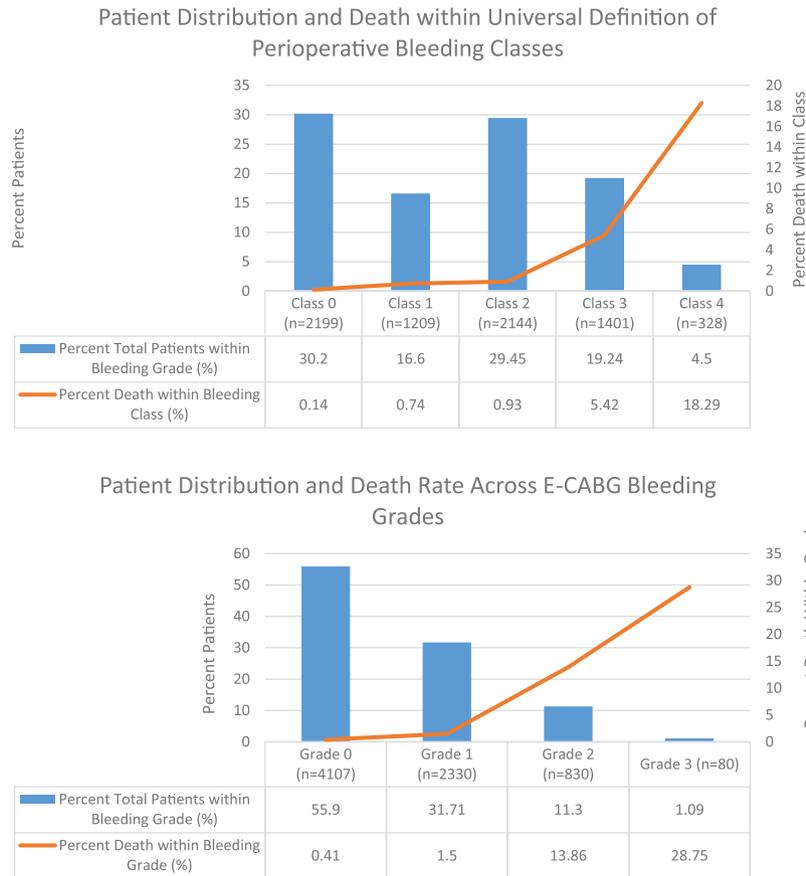
Patient Distribution and Death within Universal Definition of Perioperative Bleeding Classes

	Class 0 (n=2199)	Class 1 (n=1209)	Class 2 (n=2144)	Class 3 (n=1401)	Class 4 (n=328)
Percent Total Patients within Bleeding Grade (%)	30.2	16.6	29.45	19.24	4.5
Percent Death within Bleeding Class (%)	0.14	0.74	0.93	5.42	18.29

Patient Distribution and Death Rate Across E-CABG Bleeding Grades

	Grade 0 (n=4107)	Grade 1 (n=2330)	Grade 2 (n=830)	Grade 3 (n=80)
Percent Total Patients within Bleeding Grade (%)	55.9	31.71	11.3	1.09
Percent Death within Bleeding Grade (%)	0.41	1.5	13.86	28.75

Fig. 2. The distribution of patients across Universal definition of perioperative bleeding classes and European Coronary Artery Bypass Graft (E-CABG) classes. The proportion of patients with different severity categories is shown for each bleeding definition. The risk of death within each category is also shown.

Table 3. Unadjusted Odds Ratios Relating Measured Scores to 28-day Mortality

Universal Definition of Perioperative Bleeding Severity Score Unadjusted Odds Ratios			
Predictor	Odds Ratio	95% CI	P Value
UDPB grade			
Class 0	Reference		
Class 1	5.41	1.46–20.0	0.01
Class 2	6.15	1.81–20.90	0.004
Class 3	41.49	13.06–131.78	< 0.0001
Class 4	158.37	49.30–508.72	< 0.0001
AUC (95% CI): 0.84 (0.81, 0.87)			
E-CABG Bleeding Severity Score Unadjusted Odds Ratios			
Predictor	Odds Ratio	95% CI	P Value
E-CABG grade			
Grade 0	Reference		
Grade 1	3.62	1.95–6.74	< 0.0001
Grade 2	38.77	22.39–67.14	< 0.0001
Grade 3	106.45	52.03–217.80	< 0.0001
AUC (95% CI): 0.85 (0.82, 0.88)			

AUC, area under the curve; E-CABG, European Coronary Artery Bypass Graft; UDPB, Universal Definition of Perioperative Bleeding in Adult Cardiac Surgery.[9]

institutional databases at two hospitals in Italy. Increasing E-CABG severity grades were shown to be independently associated with higher in-hospital mortality and composite adverse events.[11] In a separate study drawn from 3,730 patients in a multicentre prospective registry encompassing 16 centers in six countries (England, Finland, France, Germany, Italy, and Sweden), six different bleeding scores were compared against each other, including E-CABG and the Universal definition of perioperative bleeding.[20] Both E-CABG and the Universal score showed good discriminative and predictive ability, with acceptable area-under-the-curve values for prediction of mortality, stroke, acute kidney injury, and sternal wound infection.[20]

Our study offers strong evidence supporting the use of either score in a clinical trial context. Despite both being useful, we also noted some key differences. Support for the Universal score and E-CABG measuring a similar construct was offered by the higher complex and redo procedures in higher bleeding grades. In terms of content validity, the Universal score captures a wider variety of clinical events associated with bleeding when assigning patients to a given category of bleeding severity. E-CABG

may better capture the adverse impact of higher transfusion needs because it scores almost exclusively transfusion volume. The Universal score also assesses transfusion volume but may better capture coagulopathy related to massive bleeding due to the scoring of recombinant activated factor VII administration, for example.[31] Thus, although the two scores have some overlap, each captures slightly different aspects relevant to major bleeding.

For both scores, it is important to note that patients undergoing surgery with a lower total hemoglobin mass to begin with are more likely to receive red cell transfusions than those with higher hemoglobin mass. Because the amount of red cell transfusions is a major component of both scores, investigators should control for patients' baseline hemoglobin mass if using these scores as endpoints. This also applies to the use of other product transfusions, such as platelets or factor replacement, which may be administered to correct preexisting deficiencies rather than in reaction to bleeding events.[9]

Despite E-CABG including fewer assessed domains than the Universal score, they both have evidence of construct and criterion validity. Interestingly, the scored domain of chest tube output does not appear to contribute significantly to the criterion validity of the Universal score, and any future revision to the Universal score may consider omitting this component. Although chest tube output has been associated with mortality in other studies, this has been in samples where patients receiving transfusion or having other significant clinical events such as reoperation associated with ongoing bleeding were not excluded.[2,32] In our study, chest tube output alone, without the presence of any other items indicative of bleeding, was not associated with increased mortality. Patients with chest tube output and other items indicating bleeding did have increased mortality, which is largely consistent with the existing literature.

The time frame for the scoring of the Universal definition of perioperative bleeding and E-CABG is notably different. E-CABG is scored based on data collected from the initial surgery throughout the entire hospital stay of the patient, whereas the assessment of the Universal score is limited to events that occur from surgery to the first perioperative day. Because E-CABG scores primarily transfusion volume, events during the hospital stay distant or unrelated to the original surgery that result in transfusion are captured and scored as bleeding, which may erroneously be associated with the original surgery. It is important to note this important distinction, which may impact on the construct validity of E-CABG.

Generally, during the conduct of a trial, the more complex the endpoints and the greater the volume of data collected, the higher the risk of incomplete data. Difficulty collecting data to assess either the Universal score or E-CABG was not a problem in the TACS study. Some investigators may prefer to use E-CABG based on the lower number of domains scored and lower burden of data collection. However, any

potential savings in data collection when using E-CABG as an outcome have to be weighed against the reduced number of potential bleeding events captured by E-CABG, particularly more severe bleeding. The probability of more severe bleeding score by the Universal score (class 3 or 4) was nearly double that of the E-CABG (grade 2 or 3) in our sample (23.75% by the Universal score *vs.* 12.39% by E-CABG). A greater proportion of patients are scored as having bleeding events, including more severe bleeding, when the Universal score is used. In addition, the distribution of patients across bleeding categories is more even with the Universal score.

Although we attempted to provide a thorough comparison of these two consensus-based scores, there are limitations to our work. In assessing criterion validity, our outcome was 28-day in-hospital all-cause mortality. This was the data we had available to us, but assessing all-cause mortality not limited to the hospital setting is important. Although limiting mortality to within 28-days allowed us to focus on the immediate perioperative period, certainly mortality beyond this time frame is important. For example, complications of severe bleeding events include acute kidney injury, which may have an impact on patient mortality beyond 28 days and would not have been captured in our data set.[33] Furthermore, in our multivariable logistic regression models examining the association of each bleeding score with mortality, we included as covariates variables that were identified as prognostically important in the literature. Our data set did not include important items scored by EuroSCORE II, a model used worldwide for the prediction of cardiac surgical risk.[34] This precluded us from including it as a covariate in our models.

The adoption of clinically sensible, consistent endpoints in cardiac surgery clinical trials evaluating bleeding has many advantages. In this study, we add to existing evidence supporting the use of these two consensus-based bleeding scores in cardiac surgery. Both the Universal definition of perioperative bleeding and E-CABG demonstrate evidence of criterion, construct, and content validity. However, the Universal score seems to capture more bleeding events and has a more uniform distribution of bleeding categories. This has implications for clinical trial sample size calculations and suggests that fewer patients may be required to demonstrate a difference between groups if the Universal score is used as an endpoint. The Universal score captures a variety of clinically important events related to bleeding beyond transfusion in the time period immediately related to surgery. E-CABG primarily captures transfusion volume over the entire length of hospital stay, which may not always be related to bleeding events associated with the original surgery. On the other hand, the individual components of E-CABG are easily collected from most hospital administrative databases and may represent significantly less burden of work for trial organizers needing to be efficient with their resources. The ongoing growth of clinical trials in cardiac surgery will benefit from the use of either scoring system in future clinical trials measuring bleeding as an outcome of interest.

Acknowledgments

Research Support

Competing Interests

Correspondence

Address correspondence to Dr. Karkouti: Toronto General Hospital, 200 Elizabeth Street, 3EN, Toronto, Ontario M5G 2C4, Canada. keyvan.karkouti@uhn.ca. Information on purchasing reprints may be found at www.anesthesiology.org or on the masthead page at the beginning of this issue. ANESTHESIOLOGY's articles are made freely accessible to all readers, for personal use only, 6 months from the cover date of the issue.

References

1. Karkouti K, O'Farrell R, Yau TM, Beattie WS; Reducing Bleeding in Cardiac Surgery Research Group: Prediction of massive blood transfusion in cardiac surgery. Can J Anaesth 2006; 53:781–94

2. Colson PH, Gaudard P, Fellahi JL, Bertet H, Faucanie M, Amour J, Blanloeil Y, Lanquetot H, Ouattara A, Picot MC; ARCOTHOVA group: Active bleeding after cardiac surgery: A prospective observational multicenter study. PLoS One 2016; 11:e0162396

3. Schulman S, Angerås U, Bergqvist D, Eriksson B, Lassen MR, Fisher W; Subcommittee on Control of Anticoagulation of the Scientific and Standardization Committee of the International Society on Thrombosis and Haemostasis: Definition of major bleeding in clinical investigations of antihemostatic medicinal products in surgical patients. J Thromb Haemost 2010; 8:202–4

4. Mehran R, Rao SV, Bhatt DL, Gibson CM, Caixeta A, Eikelboom J, Kaul S, Wiviott SD, Menon V, Nikolsky E, Serebruany V, Valgimigli M, Vranckx P, Taggart D, Sabik JF, Cutlip DE, Krucoff MW, Ohman EM, Steg PG, White H: Standardized bleeding definitions for cardiovascular clinical trials: A consensus report from the Bleeding Academic Research Consortium. Circulation 2011; 123:2736–47

5. Myles PS, Grocott MP, Boney O, Moonesinghe SR; COMPAC-StEP Group: Standardizing end points in perioperative trials: Towards a core and extended outcome set. Br J Anaesth 2016; 116:586–9

6. Prinsen CA, Vohra S, Rose MR, King-Jones S, Ishaque S, Bhaloo Z, Adams D, Terwee CB: Core Outcome Measures in

7. Koroshetz W: A core set of trial outcomes for every medical discipline? BMJ 2015; 350:h85

8. Williamson PR, Altman DG, Blazeby JM, Clarke M, Devane D, Gargon E, Tugwell P: Developing core outcome sets for clinical trials: Issues to consider. Trials 2012; 13:132

9. Dyke C, Aronson S, Dietrich W, Hofmann A, Karkouti K, Levi M, Murphy GJ, Sellke FW, Shore-Lesserson L, von Heymann C, Ranucci M: Universal definition of perioperative bleeding in adult cardiac surgery. J Thorac Cardiovasc Surg 2014; 147:1458–1463.e1

10. Biancari F, Ruggieri VG, Perrotti A, Svenarud P, Dalén M, Onorati F, Faggian G, Santarpino G, Maselli D, Dominici C, Nardella S, Musumeci F, Gherli R, Mariscalco G, Masala N, Rubino AS, Mignosa C, De Feo M, Della Corte A, Bancone C, Chocron S, Gatti G, Gherli T, Kinnunen EM, Juvonen T: European Multicenter Study on Coronary Artery Bypass Grafting (E-CABG registry): Study protocol for a prospective clinical registry and proposal of classification of postoperative complications. J Cardiothorac Surg 2015; 10:90

11. Mariscalco G, Gherli R, Ahmed AB, Zanobini M, Maselli D, Dalén M, Piffaretti G, Cappabianca G, Beghi C, Biancari F: Validation of the European Multicenter Study on Coronary Artery Bypass Grafting (E-CABG) bleeding severity definition. Ann Thorac Surg 2016; 101:1782–8

12. Higgins PG: Measuring nurses' accuracy of estimating blood loss. J Adv Nurs 1982; 7:157–62

13. Rothermel LD, Lipman JM: Estimation of blood loss is inaccurate and unreliable. Surgery 2016; 160:946–53

14. Vranckx P, Leonardi S, Tebaldi M, Biscaglia S, Parrinello G, Rao SV, Mehran R, Valgimigli M: Prospective validation of the Bleeding Academic Research Consortium classification in the all-comer PRODIGY trial. Eur Heart J 2014; 35:2524–9

15. Fleming TR, DeMets DL: Surrogate end points in clinical trials: Are we being misled? Ann Intern Med 1996; 125:605–13

16. Prentice RL: Surrogate endpoints in clinical trials: Definition and operational criteria. Stat Med 1989; 8:431–40

17. Karkouti K, Callum J, Wijeysundera DN, Rao V, Crowther M, Grocott HP, Pinto R, Scales DC; TACS Investigators: Point-of-care hemostatic testing in cardiac surgery: A stepped-wedge clustered randomized controlled trial. Circulation 2016; 134:1152–62

18. Ricci Z, Cruz DN, Ronco C: Classification and staging of acute kidney injury: Beyond the RIFLE and AKIN criteria. Nat Rev Nephrol 2011; 7:201–8

19. Kinnunen EM, Juvonen T, Airaksinen KE, Heikkinen J, Kettunen U, Mariscalco G, Biancari F: Clinical significance and determinants of the universal definition of perioperative bleeding classification in patients undergoing coronary artery bypass surgery. J Thorac Cardiovasc Surg 2014; 148:1640–1646.e2

20. Brascia D, Reichart D, Onorati F, Perrotti A, Ruggieri VG, Bounader K, Verhoye JP, Santarpino G, Fischlein T, Maselli D, Dominici C, Mariscalco G, Gherli R, Rubino AS, De Feo M, Bancone C, Gatti G, Santini F, Dalén M, Saccocci M, Faggian G, Tauriainen T, Kinnunen EM, Nicolini F, Gherli T, Rosato S, Biancari F: Validation of bleeding classifications in coronary artery bypass grafting. Am J Cardiol 2017; 119:727–33

21. Nashef SA, Roques F, Sharples LD, Nilsson J, Smith C, Goldstone AR, Lockowandt U: EuroSCORE II. Eur J Cardiothorac Surg 2012; 41:734–45

22. Ivanov J, Borger MA, Rao V, David TE: The Toronto Risk Score for adverse events following cardiac surgery. Can J Cardiol 2006; 22:221–7

23. Greiff G, Pleym H, Stenseth R, Berg KS, Wahba A, Videm V: Prediction of bleeding after cardiac surgery: Comparison

of model performances: A prospective observational study. J Cardiothorac Vasc Anesth 2015; 29:311–9

24. Dupuis JY, Wang F, Nathan H, Lam M, Grimes S, Bourke M: The cardiac anesthesia risk evaluation score: A clinically useful predictor of mortality and morbidity after cardiac surgery. ANESTHESIOLOGY 2001; 94:194–204

25. Aktuerk D, McNulty D, Ray D, Begaj I, Howell N, Freemantle N, Pagano D: National administrative data produces an accurate and stable risk prediction model for short-term and 1-year mortality following cardiac surgery. Int J Cardiol 2016; 203:196–203

26. Tran DT, Dupuis JY, Mesana T, Ruel M, Nathan HJ: Comparison of the EuroSCORE and Cardiac Anesthesia Risk Evaluation (CARE) score for risk-adjusted mortality analysis in cardiac surgery. Eur J Cardiothorac Surg 2012; 41:307–13

27. Rankin JS, He X, O'Brien SM, Jacobs JP, Welke KF, Filardo G, Shahian DM: The Society of Thoracic Surgeons risk model for operative mortality after multiple valve surgery. Ann Thorac Surg 2013; 95:1484–90

28. O'Brien SM, Shahian DM, Filardo G, Ferraris VA, Haan CK, Rich JB, Normand SL, DeLong ER, Shewan CM, Dokholyan RS, Peterson ED, Edwards FH, Anderson RP: The Society of Thoracic Surgeons 2008 cardiac surgery risk models: Part 2. Isolated valve surgery. Ann Thorac Surg 2009; 88:S23–42

29. Shahian DM, O'Brien SM, Filardo G, Ferraris VA, Haan CK, Rich JB, Normand SL, DeLong ER, Shewan CM, Dokholyan RS, Peterson ED, Edwards FH, Anderson RP: The Society of Thoracic Surgeons 2008 cardiac surgery risk models: Part 3. Valve plus coronary artery bypass grafting surgery. Ann Thorac Surg 2009; 88:S43–62

30. Cockcroft DW, Gault MH: Prediction of creatinine clearance from serum creatinine. Nephron 1976; 16:31–41

31. Karkouti K, Beattie WS, Arellano R, Aye T, Bussieres JS, Callum JL, Cheng D, Heinrich L, Kent B, Lee TW, MacAdams C, Mazer CD, Muirhead B, Rochon AG, Rubens FD, Sawchuk C, Wang S, Waters T, Wong BI, Yau TM: Comprehensive Canadian review of the off-label use of recombinant activated factor VII in cardiac surgery. Circulation 2008; 118:331–8

32. Christensen MC, Dziewior F, Kempel A, von Heymann C: Increased chest tube drainage is independently associated with adverse outcome after cardiac surgery. J Cardiothorac Vasc Anesth 2012; 26:46–51

33. O'Connor ME, Hewson RW, Kirwan CJ, Ackland GL, Pearse RM, Prowle JR: Acute kidney injury and mortality 1 year after major non-cardiac surgery. Br J Surg 2017; 104:868–76

34. Noyez L, Kievit PC, van Swieten HA, de Boer MJ: Cardiac operative risk evaluation: The EuroSCORE II: Does it make a real difference? Neth Heart J 2012; 20:494–8

Appendix

The Transfusion Avoidance in Cardiac Surgery (TACS Research Group) included:

Keyvan Karkouti, M.D.
Jeannie Callum, M.D.
Duminda N. Wijeysundera, M.D., Ph.D.
Vivek Rao, M.D., Ph.D.
Mark Crowther, M.D.
Hilary P. Grocott, M.D.
Ruxandra Pinto, Ph.D.
Damon C. Scales, M.D., Ph.D.

Blaine Achen, M.D.
Sukhpal Brar, M.D.
Doug Morrison, M.D.
David Wong, M.D.
Jean S. Bussières, M.D.
Tonya de Waal, M.D.
Christopher Harle, M.D.
Étienne de Médicis, M.D., M.Sc.
Charles McAdams, M.D.
Summer Syed, M.D.
Diem Tran, M.D.
Terry Waters, M.D.