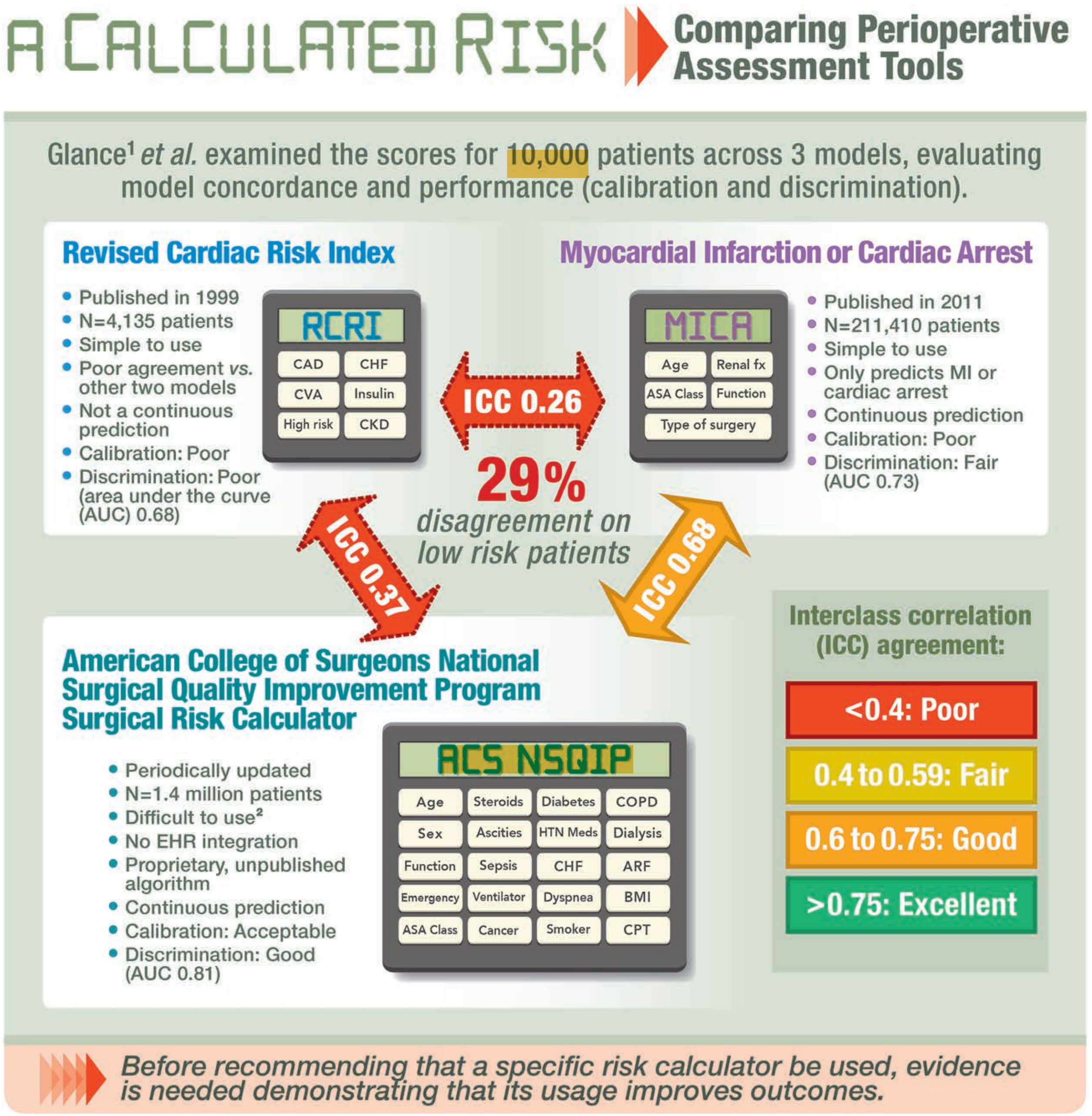
ANESTHESIOLOGY



ACS, American College of Surgeons; ARF, acute renal failure; ASA, American Society of Anesthesiologists; Ascites, ascites within 30 days prior to surgery; BMI, body mass index; CAD, coronary artery disease; Cancer, disseminated cancer; CHF, congestive heart failure in 30 days prior to surgery; CKD, chronic kidney disease; COPD, chronic obstructive pulmonary disease, severe; CPT, Current Procedural Terminology, specific code; CVA, cardiovascular accident; EHR, electronic health record; Function, functional status; High risk, high-risk surgical procedure; HTN meds, hypertension requiring medication; Insulin, diabetes requiring the usage of insulin; MI, myocardial infarction; MICA, Myocardial Infarction or Cardiac Arrest; NSQIP, National Surgical Quality Improvement Program; RCRI, Revised Cardiac Risk Index; Renal fx, renal function; Sepsis, sepsis within 48 hr prior to surgery; Smoker, current smoker within 1 yr; Ventilator, ventilator dependent.

Infographic created by Jonathan P. Wanderer, Vanderbilt University Medical Center, and James P. Rathmell, Brigham and Women's Health Care/Harvard Medical School. Illustration by Annemarie Johnson, Vivo Visuals. Address correspondence to Dr. Wanderer: jonathan.p.wanderer@vanderbilt.edu.

1. Glance LG, Faden E, Dutton RP, Lustik SJ, Li Y, Eaton MP, Dick AW. Impact of the choice of risk model for identifying low-risk patients using the 2014 American College of Cardiology/American Heart Association perioperative guidelines. ANESTHESIOLOGY 2018; 129:889–900

2. Fleisher LA. Preoperative cardiac evaluation before noncardiac surgery: Reverend Bayes's risk indices and optimal variables. ANESTHESIOLOGY 2018; 129:867–8

Impact of the Choice of Risk Model for Identifying Low-risk Patients Using the 2014 American College of Cardiology/American Heart Association Perioperative Guidelines

Laurent G. Glance, M.D., Eric Faden, M.D., Richard P. Dutton, M.D., M.B.A., Stewart J. Lustik, M.D., M.B.A., Yue Li, Ph.D., Michael P. Eaton, M.D., Andrew W. Dick, Ph.D.

ABSTRACT

Background: The 2014 American College of Cardiology Perioperative Guideline recommends risk stratifying patients scheduled to undergo noncardiac surgery using either: (1) the Revised Cardiac Index; (2) the American College of Surgeons National Surgical Quality Improvement Program Surgical Risk Calculator; or (3) the Myocardial Infarction or Cardiac Arrest calculator. The aim of this study is to determine how often these three risk-prediction tools agree on the classification of patients as low risk (less than 1%) of major adverse cardiac event.

Methods: This is a retrospective observational study using a sample of 10,000 patient records. The risk of cardiac complications was calculated for the Revised Cardiac Index and the Myocardial Infarction or Cardiac Arrest models using published coefficients, and for the American College of Surgeons National Surgical Quality Improvement Program Surgical Risk Calculator using the publicly available website. The authors used the intraclass correlation coefficient and kappa analysis to quantify the degree of agreement between these three risk-prediction tools.

Results: There is good agreement between the American College of Surgeons National Surgical Quality Improvement Program and Myocardial Infarction or Cardiac Arrest estimates of major adverse cardiac events (intraclass correlation coefficient = 0.68, 95% CI: 0.66 to 0.70), while only poor agreement between (1) American College of Surgeons National Surgical Quality Improvement Program Surgical Risk Calculator and the Revised Cardiac Index (intraclass correlation coefficient = 0.37; 95% CI: 0.34 to 0.40), and (2) Myocardial Infarction or Cardiac Arrest and Revised Cardiac Index (intraclass correlation coefficient = 0.26; 95% CI: 0.23 to 0.30). The three prediction models disagreed 29% of the time on which patients were low risk.

Conclusions: There is wide variability in the predicted risk of cardiac complications using different risk-prediction tools. Including more than one prediction tool in clinical guidelines could lead to differences in decision-making for some patients depending on which risk calculator is used. (ANESTHESIOLOGY 2018; 129:889-900)

MORE than 200 million major surgical procedures are performed worldwide each year.[1] Eight percent of patients undergoing major noncardiac surgery experience myocardial injury.[2] Perioperative mortality is the third leading cause of death in the United States,[3] and cardiac complications cause one third of all perioperative deaths.[4] Current clinical practice guidelines endorsed by the American College of Cardiology and the American Heart Association recommends a stepwise approach to the cardiac assessment of patients scheduled to undergo noncardiac surgery. According to these guidelines, patients whose estimated perioperative risk of major adverse cardiac event is less than 1% can proceed to surgery without undergoing further cardiac testing.[5] The use of a validated risk-prediction tool to predict the risk of perioperative major adverse cardiac event is a Class IIa recommendation.[5] Three risk models are recommended by the American College of Cardiology/American Heart Association guidelines

Editor's Perspective

What We Already Know about This Topic

- The Revised Cardiac Risk Index, and the risk calculators based on the National Surgical Quality Improvement Program can be used to assess the risk of cardiac adverse events after noncardiac surgery
- Recent clinical practice guidelines recommend the use of one of these calculators
- The agreement across these calculators is poorly understood and has not been robustly tested in a single analysis

What This Article Tells Us That Is New

- Thirty percent of predictions regarding high *versus* low risk are discordant across the risk calculators
- The choice of risk-prediction tool could have an impact on the calculated risk and subsequent clinical decisions

This article is featured in "This Month in Anesthesiology," page 1A. Corresponding article on page 867. Supplemental Digital Content is available for this article. Direct URL citations appear in the printed text and are available in both the HTML and PDF versions of this article. Links to the digital files are provided in the HTML text of this article on the Journal's Web site (www.anesthesiology.org). This article has an audio podcast. This article has a visual abstract available in the online version.

Submitted for publication July 7, 2017. Accepted for publication June 5, 2018. From the Department of Anesthesiology (L.G.G., E.F., S.J.L., M.P.E.) and the Department of Public Health Sciences (L.G.G., Y.L.), University of Rochester School of Medicine, Rochester, New York; RAND Health, Boston, Massachusetts (L.G.G., A.W.D.); and U.S. Anesthesia Partners, Dallas, Texas (R.P.D.).

for predicting the risk of major adverse cardiac event: (1) the Revised Cardiac Risk Index[6]; (2) The American College of Surgeons National Surgical Quality Improvement Program Surgical Risk Calculator;[7] and (3) the National Surgical Quality Improvement Program Myocardial Infarction or Cardiac Arrest (myocardial infarction or cardiac arrest) calculator.[8]

Risk stratification, using one of these three prediction models to separate *low-risk* (less than 1%) from *elevated-risk* (greater than or equal to 1%) patients, is a key decision point in the American College of Cardiology/American Heart Association perioperative guideline. In theory, prediction models allow clinicians to more accurately estimate risk rather than relying on clinical judgement alone. However, while drugs and therapies must undergo U.S. Food and Drug Administration approval prior to clinical use, the decision to incorporate clinical prediction models in patient care is based on peer-review alone. The American College of Cardiology/American Heart Association guideline does not specify which model to use for risk stratification. Twenty years ago, Iezzoni published a seminal paper *The Risks of Risk Adjustment* showing that hospital quality is partly a function of the choice of the risk adjustment model.[9] Our goal in this exploratory study is to determine how often the three prediction models recommended in the American College of Cardiology/American Heart Association perioperative guidelines agree on the classification of patients scheduled to undergo major noncardiac surgery as low risk *versus* elevated risk. We will use data from the American College of Surgeons National Surgical Quality Improvement Program database to calculate the predicted risk of major adverse cardiac event for individual patients using each of these three prediction models. We expect to find that predicted patient risk is also, in part, a function of the choice of prediction models. Our findings may help inform the process of developing best practices for incorporating prediction models in clinical practice guidelines.

Materials and Methods

Data Source

This analysis was based on data from the American College of Surgeons National Surgical Quality Improvement Program participant use data file for patients undergoing noncardiac surgery in 2012. The database includes comprehensive clinical information on patient demographics, preoperative risk factors, and 30-day postoperative mortality and complications.[10] We used the National Surgical Quality Improvement Program definition of major adverse cardiac event as cardiac arrest or myocardial infarction (appendix).[7] Myocardial infarction is defined by the National Surgical Quality Improvement Program as either (1) "documentation of electrocardiogram changes indicative of acute myocardial infarction (ST elevation > 1 mm in two contiguous leads, new left bundle branch block, or new Q waves in two or more contiguous leads)"; (2) "new elevation of troponin greater than three times the upper level of the reference range in the setting of suspected myocardial ischemia"; or (3) "physician diagnosis of myocardial infarction.[10]" These data elements are prospectively collected by trained Surgical Clinical Reviewers using chart abstraction and through extraction from hospitals' information systems.[10] Data quality is verified using auditing comparing chart data with data submitted by participating sites. Systematic case sampling is used to avoid bias in case selection.[10]

Study Population

We first identified 71,575 patients in the American College of Surgeons National Surgical Quality Improvement Program participant use data file who underwent one of the major noncardiac surgical procedures tracked by the Centers for Disease Control and Prevention National Healthcare Safety Network, excluding very low-risk procedures (*e.g.*, appendectomy, breast procedures, pacemakers), and cardiac cases. We excluded cases from hospitals submitting a minimal data set which did not include data elements for history of myocardial infarction and stroke; these data elements are necessary to calculate the risk of major adverse cardiac event using the Revised Cardiac Risk Index (fig. 1). Procedures were then divided into high-risk or low/intermediate-risk surgeries based on the approach used in the derivation of the Revised Cardiac Risk Index [6] and the procedure stratification described in the American College of Cardiology/American Heart Association 2007 guidelines.[11] We then constructed a stratified sample of 10,000 records in which we over-sampled patients undergoing high-risk procedures because we assumed that surgeons are more likely to obtain cardiology consultation for higher-risk surgeries. To construct our analytic data set, we excluded American Society of Anesthesiologists (ASA) physical status V and emergency cases (emergency cases do not undergo further cardiac testing in the American College of Cardiology/American Heart Association guidelines), records with missing ASA physical status or functional status, and records missing American College of Surgeons National Surgical Quality Improvement Program risk of major adverse cardiac event (fig. 1). For the purpose of our analysis, we assumed that patients with no recorded value of serum creatinine had normal renal function, unless they were coded as having renal failure or dialysis (n = 880). *A priori*, we limited the size of the analytic data set to 10,000 observations so that the number of inquiries submitted to the web-based risk calculator did not interfere with the ability of clinicians to use the American College of Surgeons risk calculator for clinical purposes. We chose to reduce our sample size by including about 10% of the intermediate- or low-risk surgery cases, while retaining most of the high-risk procedures. This decision was driven by our assumption that the American College of Cardiology/American Heart Association algorithm is more often used by clinicians for patients undergoing higher-risk surgeries, and that our findings would thus have more clinical relevance when applied to a population that reflected the clinical use of the American College of Cardiology/

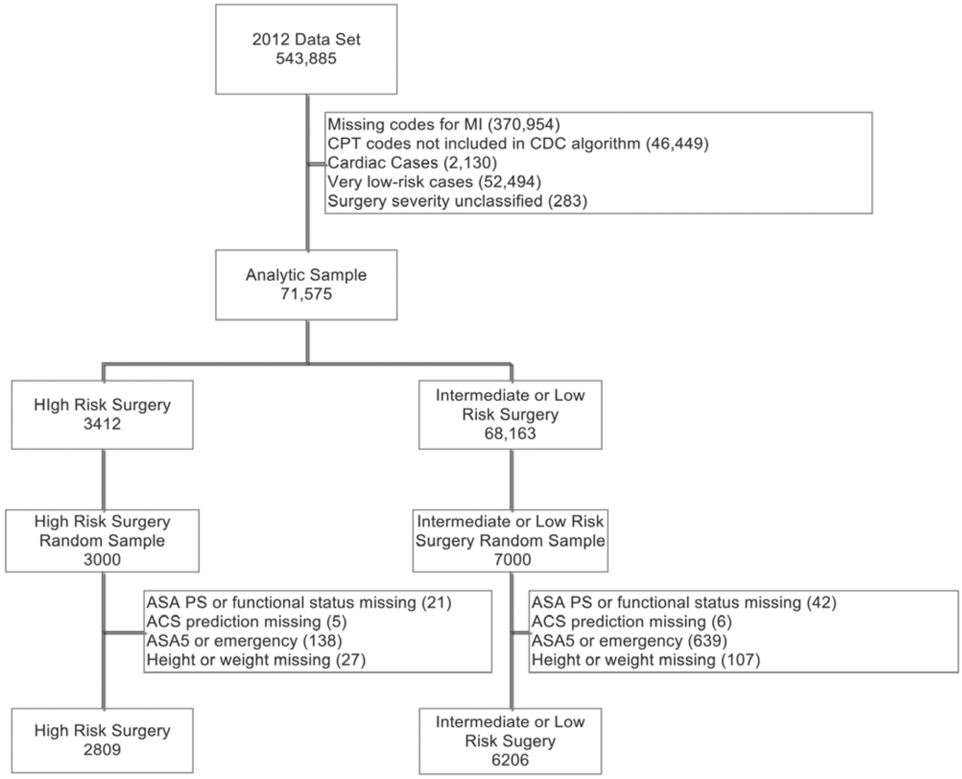


Fig. 1. Flow diagram illustrating selection of patients included in the analytic cohort. “High-risk surgery” and “intermediate- or low-risk surgery” refers to risk of surgery, and does not take into account patient risk. Procedures were classified as high-risk or low/intermediate-risk surgeries based on the approach used in the derivation of the Revised Cardiac Risk Index and the procedure stratification described in the American College of Cardiology/American Heart Association 2007 guidelines. ACS = American College of Surgeons; ASA PS = American Society of Anesthesiologists physical status; MI = myocardial infarction.

American Heart Assocation guideline. We did, however, perform a sensitivity analysis (described in “Statistical Analysis”) to determine whether our findings were altered by our sampling strategy.

Statistical Analysis

The risk of major adverse cardiac event was calculated for the Revised Cardiac Risk Index and the National Surgical Quality Improvement Program Myocardial Infarction or Cardiac Arrest models using published coefficients.[8,12] We submitted the records in our analytic data set to the publicly available American College of Surgeons National Surgical Quality Improvement Program Surgical Risk Calculator website to obtain risk estimates based on the American College of Surgeons National Surgical Quality Improvement Program prediction model.[13] The risk of major adverse cardiac event predicted by the American College of Surgeons National Surgical Quality Improvement Program and the National Surgical Quality Improvement Program Myocardial Infarction or Cardiac Arrest prediction tools are continuous measures, whereas the Revised Cardiac Risk Index risk of major adverse cardiac event can only take on one of four discrete values: 0.4, 0.9, 6.6, and 11%.

The goal of this study was to examine the degree of agreement among these three risk-prediction tools. We performed two analyses. In the first analysis, we quantified the agreement in the patient-level risk estimates of major adverse cardiac event between each pair of the three risk-prediction tools using the intraclass correlation coefficient: (1) Revised Cardiac Risk Index *versus* National Surgical Quality Improvement Program Myocardial Infarction or Cardiac Arrest; (2) Revised Cardiac Risk Index *versus* American College of Surgeons National Surgical Quality Improvement Program; and (3) National Surgical Quality Improvement Program Myocardial Infarction or Cardiac Arrest *versus* American College of Surgeons National Surgical Quality Improvement Program. The intraclass correlation coefficient measures the interrater reliability and reflects the agreement between two risk-prediction scores (*i.e.*, Revised Cardiac Risk Index *vs.* American College of Surgeons National Surgical Quality Improvement Program) on the risk of major adverse cardiac event for the same sample of patients. The intraclass correlation coefficient is based on a two-way mixed effects model, where the dependent variable y is the predicted risk of major adverse cardiac event, the risk score S is an indicator variable specifying the risk score (*i.e.*, Revised Cardiac Risk Index *vs.* National Surgical Quality Improvement Program Myocardial Infarction or Cardiac Arrest), patient is specified as a random effect P, and ε is the error term:

$$y = \text{constant} + \text{S} + \text{P} + \varepsilon$$

The intraclass correlation coefficient is calculated using[14]:

$$\text{ICC} = \frac{\sigma_p^2}{\sigma_p^2 + \sigma_S^2 + \sigma_\varepsilon^2}$$

where σ_p^2 is the patient variance, σ_S^2 is the variance of the fixed effect indicating which score is being used, and σ_ε^2 is the variance of the error term.

If the scores (*e.g.*, Revised Cardiac Risk Index and American College of Surgeons National Surgical Quality Improvement Program) are in close agreement on the value of the of the calculated patient risk, then $\sigma_p^2 \gg \sigma_S^2$ and $\sigma_p^2 \gg \sigma_\varepsilon^2$, and the intraclass correlation coefficient will be close to 1. Essentially, the intraclass correlation coefficient represents the signal-to-noise ratio.[15] The closer the intraclass correlation coefficient is to "1," the stronger the level of agreement between the 2 score. Intraclass correlation coefficient values less than 0.4 suggest poor reliability, between 0.4 to 0.59 represent fair reliability, between 0.60 and 0.74 good, and between 0.75 and 1.00 excellent reliability.[16]

Since the key decision point in the American College of Cardiology/American Heart Association algorithm is to separate patients into low-risk (less than 1%) and high-risk (greater than or equal to 1%) groups, we performed a second analysis in which we examined the extent to which the three risk-prediction tools agreed on the classification of patients into low- and high-risk groups. Patient records were first dichotomized into low risk of major adverse cardiac event (less than 1%) *versus* high risk of major adverse cardiac event (greater than or equal to 1%), and then kappa analysis was used to examine the level of agreement for all three risk models. Bootstrapping was used to construct 95% CI around the kappa statistic. Using the kappa statistic, values less than 0 suggest poor agreement, 0.00 to 0.20 slight, 0.21 to 0.40 fair, 0.41 to 0.60 moderate, 0.61 to 0.80 substantial, and 0.81 to 1.00 almost perfect agreement.[17]

We performed three sensitivity analyses. First, we repeated our analyses after excluding observations with missing creatinine values: complete case analysis. Second, we performed weighted kappa analysis to examine the agreement between the three risk-prediction tools, using the STATA-defined weight *w*. In the third sensitivity analysis, we over-sampled observations for intermediate- and low-risk surgeries to create an analytic sample that more closely resembles the ratio of high-risk to intermediate/low-risk surgeries (1:20) in original analytic sample (fig. 1). We created 10 bootstrap samples of the observations in the intermediate/low-risk surgery patients using sampling with replacement (62,060) and appended this to the sample of high-risk surgery patients (2,809). We then used the intraclass correlation coefficient and kappa analysis to examine the agreement between the three risk-prediction tools in this new analytic data set.

The goal of this exploratory study was not to perform an independent validation of these three risk-prediction tools. We did, however, perform a secondary analysis in which we examined the performance of each of these models using measures of discrimination (C statistic) and calibration (calibration curve, Hosmer–Lemeshow statistic). The C statistic is the probability that a patient selected at random who experienced a major adverse cardiac event had a higher predicted risk of a major adverse cardiac event compared to a randomly selected patient who did not experience a major adverse cardiac event.[17] If the C statistic is 0.5, then the risk-prediction tool is no better than the flip of a coin. Higher C statistics indicates that a risk model is better able to discriminate between patients who experienced an event (*i.e.*, major adverse cardiac event) and those that did not.[18] The Hosmer–Lemeshow statistic assesses goodness-of-fit by first ranking the observations by the predicted probability of the event (major adverse cardiac event) into deciles of risk, and then comparing the observed and the predicted number of events within each decile using a chi-square–like statistic. *P* values greater than 0.05 for the Hosmer–Lemeshow statistic indicate acceptable model calibration.[18] Since the Revised Cardiac Risk Index definition of major adverse cardiac event is somewhat different than the American College of Surgeons National Surgical Quality Improvement Program and the National Surgical Quality Improvement Program Myocardial Infarction or Cardiac Arrest definition, and not all of the components of the Revised Cardiac Risk Index definition of major adverse cardiac event major adverse cardiac event are included in the National Surgical Quality Improvement Program data, these comparisons must be interpreted with caution (appendix). These differences in the definition of major adverse cardiac event do not affect the validity of the primary analyses since the American College of Cardiology/American Heart Association guidelines assume that each of these three recommended prediction tools predicts the same clinical outcome.

We would also like to highlight the fact that we compared the Revised Cardiac Risk Index risk of major adverse cardiac event (which has only four possible values) to models (American College of Surgeons National Surgical Quality Improvement Program and National Surgical Quality Improvement Program Myocardial Infarction or Cardiac Arrest) which allow nearly continuous measures of predicted risk. By design, Lee *et al.*[6] converted their multivariable risk prediction model to a risk index to make it easier for clinicians to use in daily practice. However, converting a regression model to a risk index (*i.e.*, Revised Cardiac Risk Index) reduces the accuracy of predictions. It is possible that the extent of agreement between the predictions based on the model used to create the Revised Cardiac Risk Index and the other two risk-prediction tools would have been higher than that with the Revised Cardiac Risk Index. Nonetheless, we decided *a priori* to examine the extent of agreement between the Revised Cardiac Risk Index predictions, as opposed to model-based predictions, to the two other risk-prediction tools because the Revised Cardiac Risk Index is used as a risk index in clinical practice. Of note, the full set of model coefficients necessary to calculate the model-based predictions for the Revised Cardiac Risk Index model are not published, only the risk index is available.

The main analyses used in this study were specified prior to initiating our analyses. We modified our original analytic plan to use the intraclass correlation coefficient (instead of linear regression analysis) to examine the agreement between the risk-prediction tools. Our original protocol did not specify that the analytic set would be limited to about 10,000 observations in order to limit the demands placed on the National Surgical Quality Improvement Program calculator website.

Data management and statistical analyses were performed using STATA SE/MP Version 14.2 (Stata Corp, USA). All statistical tests were two-tailed and P values less than 0.05 were considered significant. This study was determined to be exempt from Institutional Review Board review (RSRB00066296).

Results

Descriptive Statistics

Table 1 summarizes patient demographics. Sixty percent of the patients were under 65 yr of age and 97% were functionally independent. Over half of the patients were ASA physical status III or IV. Six percent of the patients had a history of diabetes requiring insulin and 6.5% had a history of cerebrovascular disease. The percentage of patients with a history of ischemic heart disease and congestive heart failure was 0.8% and 0.5%, respectively. One percent of the patients experienced major adverse cardiac event.

Agreement between Risk Indices

Figure 2 displays box plots comparing risk prediction using: (1) Revised Cardiac Risk Index *versus* American College of Surgeons National Surgical Quality Improvement Program risk calculator; and (2) Revised Cardiac Risk Index *versus* National Surgical Quality Improvement Program Myocardial Infarction or Cardiac Arrest risk calculator. Figure 3 displays scatter plots of: (1) Revised Cardiac Risk Index *versus* American College of Surgeons National Surgical Quality Improvement Program risk calculator; (2) Revised Cardiac Risk Index *versus* National Surgical Quality Improvement Program Myocardial Infarction or Cardiac Arrest risk calculator; and (3) American College of Surgeons National Surgical Quality Improvement Program risk calculator *versus* National Surgical Quality Improvement Program Myocardial Infarction or Cardiac Arrest risk calculator. There is good agreement between American College of Surgeons National Surgical Quality Improvement Program and National Surgical Quality Improvement Program Myocardial Infarction or Cardiac Arrest estimates of major adverse cardiac event (intraclass correlation coefficient = 0.68, 95% CI: 0.66, 0.70), and poor agreement between (1) American College of Surgeons National Surgical Quality Improvement Program and Revised Cardiac Risk Index (intraclass correlation coefficient = 0.37; 95% CI: 0.34, 0.40), and (2) National Surgical Quality Improvement Program Myocardial Infarction or Cardiac Arrest and Revised Cardiac Risk Index (intraclass correlation coefficient = 0.26; 95% CI: 0.23, 0.30). We found similar values for the intraclass correlation coefficient using complete case analysis (table 2). In the second sensitivity analysis in which we over-sampled intermediate- and low-risk surgeries, we found slightly worse agreement (table 2).

The three prediction models disagreed 29% of the time in which patients were low risk (29.2%; 95% CI: 28.1 to 30.0; table 3). Kappa analysis demonstrated fair agreement between the three scores (kappa 0.34; 95% CI: 0.33 to 0.36; table 3). The American College of Surgeons National Surgical Quality Improvement Program risk calculator exhibited fair agreement with the Revised Cardiac Risk Index (kappa = 0.22; 95% CI: 0.20, 0.25), while the National Surgical Quality Improvement Program Myocardial Infarction or Cardiac Arrest risk calculator exhibited slight agreement with the Revised Cardiac Risk Index (kappa = 0.18; 95% CI: 0.16, 0.20; tables 3 and 4). We obtained similar results when we performed a weighted kappa analysis based on the original analytic sample, and when we performed a complete case analysis (table 2). We found worse agreement when we over-sampled intermediate- and low-risk surgeries (table 2 and Supplemental Digital Content table, http://links.lww.com/ALN/B758).

Model Performance

In our secondary analysis, we found that the American College of Surgeons risk calculator exhibited the best discrimination (C statistic 0.81; 95% CI: 0.77, 0.85) compared to both the Revised Cardiac Risk Index (C statistic 0.68; 95% CI: 0.62, 0.73) and the National Surgical Quality Improvement Program Myocardial Infarction or Cardiac Arrest calculator (0.73; 95% CI: 0.68, 0.77). The American College of Surgeons National Surgical Quality Improvement Program calculator exhibited acceptable calibration (Hosmer–Lemeshow statistic 17.14; $P = 0.07$), whereas the National Surgical Quality Improvement Program Myocardial Infarction or Cardiac Arrest calculator (Hosmer–Lemeshow statistic 42.1; $P < 0.001$) and Revised Cardiac Risk Index (Hosmer–Lemeshow statistic 18.6; $P < 0.001$) were not as well calibrated. Calibration graphs are shown in the Supplemental Digital Content (http://links.lww.com/ALN/B758).

Discussion

We examined the extent to which the risk-prediction tools in the 2014 American College of Cardiology/American Heart Association perioperative clinical practice guideless agreed on the results of risk stratification. The Revised Cardiac Risk Index, which is the most widely used tool to risk stratify

Table 1. Clinical Characteristics

	Patients, n	Patients, %
ACS risk factors		
Age, yr		
Under 65	5,421	60.1
65–74	2,133	23.7
75–84	1,204	13.4
85 or older	257	2.85
Female	5,123	56.8
Functional status		
Independent	8,776	97.4
Partially dependent	201	2.23
Totally dependent	38	0.42
ASA class		
I	396	4.39
II	3,515	39.0
III	4,542	50.4
IV	562	6.23
Steroid use for chronic condition	378	4.19
Ascites within 30 days before surgery	45	0.5
Systemic sepsis within 48 h before surgery		
SIRS	129	1.43
Sepsis	78	0.87
Septic shock	15	0.17
Ventilator dependent	21	0.23
Disseminated cancer	610	6.77
Diabetes		
Oral	944	10.5
Insulin	546	6.06
Hypertension requiring medication	4,585	50.9
Congestive heart failure in 30 days before surgery	43	0.48
Dyspnea		
With moderate exertion	737	8.18
At rest	49	0.54
Current smoker within 1 yr	1,713	19
History of severe chronic obstructive pulmonary disease	457	5.07
Dialysis	94	1.04
Acute renal failure	21	0.23
Height, inches (IQR)	66	63, 69
Weight, lbs (IQR)	178	149, 215
Revised Cardiac Risk Index risk factors		
High-risk surgery	2,809	31.2
History of ischemic heart disease	69	0.77
History of congestive heart failure	43	0.48
History of cerebrovascular disease	582	6.46
Preoperative treatment with insulin	554	6.06
Preoperative creatinine > 2 mg/dl	205	2.27
Centers for Disease Control surgery group		
Abdominal aortic aneurysm repair	332	3.68
Bile duct, liver, or pancreatic surgery	1887	20.9
Carotid endarterectomy	257	2.85
Cholecystectomy	934	10.4
Colon surgery	810	8.99
Craniotomy	168	1.86
Gastric surgery	726	8.05
Hip arthroplasty	446	4.95
Abdominal hysterectomy	436	4.84
Knee arthroplasty	674	7.48
Laminectomy	121	1.34
Kidney surgery	474	5.26
Ovarian surgery	153	1.7
Prostate surgery	242	2.68
Peripheral vascular surgery	190	2.11
Rectal surgery	67	0.74
Small bowel surgery	245	2.72
Spleen surgery	29	0.32
Thoracic surgery	380	4.22
Ventricular shunt	180	2
Exploratory laparotomy	264	2.93
Outcomes		
Cardiac arrest requiring cardiopulmonary resuscitation	37	0.41
Myocardial infarction	56	0.62
Composite outcome	91	1.01

ACS = American College of Surgeons; ASA = American Society of Anesthesiologists; IQR = interquartile range; SIRS = systemic inflammatory response syndrome.

patients at risk of major adverse cardiac event,[19] exhibited only fair-to-poor agreement with two newer tools created using data from the American College of Surgeons. We found that these risk models disagreed about 30% of the time on which patients were at low risk *versus* elevated risk of major adverse cardiac event. According to the American College of Cardiology/American Heart Association algorithm, patients stratified as low risk should proceed directly to surgery, and would not be considered for possible stress testing or coronary revascularization.[5] Our findings suggest that the calculated risk of major adverse cardiac event, and the need for further cardiac testing, is at least partly a function of which risk-prediction tool is used to assess patient risk in the 2014 American College of Cardiology/American Heart Association perioperative practice guideline.

The goal of predictive analytics is to help tailor diagnostic and treatment interventions to patient risk. Historically, physicians estimated patient risk using knowledge acquired during training, clinical experience, and from the shared experiences of colleagues.[20] In the past, the American College of Cardiology/American Heart Association perioperative guidelines did not include a risk-prediction tool.[21] In theory, risk-prediction tools can more precisely assess patient risk using multivariable models that can simultaneously account for multiple risk factors. Predictions, however, are a function of which patient risks are included in the prediction model, and different models will produce different estimates of patient risk. There is a large body of literature showing that the same hospitals can be classified either as high-quality or low-quality, depending on which risk adjustment model is used.[9,22] In the current study, we found that the same patient

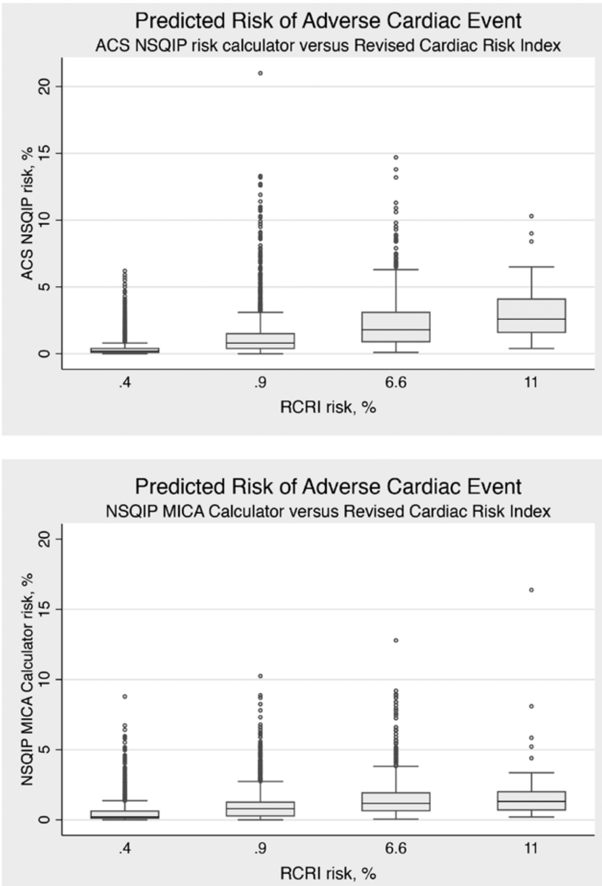


Fig. 2. Distribution of risk estimates of major adverse cardiac event as a function of: (*top*) the American College of Surgeons National Surgical Quality Improvement Program (ACS NSQIP) risk calculator *versus* the Revised Cardiac Risk Index (RCRI); and (*bottom*) the National Surgical Quality Improvement Program Myocardial Infarction or Cardiac Arrest (NSQIP MICA) calculator *versus* the RCRI. The *bottom* and *top of the box* are the first and third quartiles; the *upper whisker* is 1.5 × the width of the upper quartile, and the *lower whisker* is 1.5 × the width of the lower quartile. These figures display the upward bias of the RCRI risk predictions relative to the ACS NSQIP and the NSQIP MICA risk prediction tools for the two higher-risk RCRI categories.

could be classified as low risk or not low risk depending on which risk-prediction tool was used.

Selecting a best-in-class risk-prediction tool should take several factors into account, including face validity, statistical performance, and generalizability. Face validity is subjective, and in this case, implies that the model includes risk factors that would reasonably be associated with the risk of major adverse cardiac event (*e.g.*, cardiac risk, surgical complexity). Risk predictions are only as good as the underlying models; the statistical performance of risk models is evaluated using measures of discrimination and calibration.[23] Most models are internally validated using data from the same database that was used to create the models. External validation using a different database is more desirable because it establishes whether a risk model accurately predicts outcomes beyond the

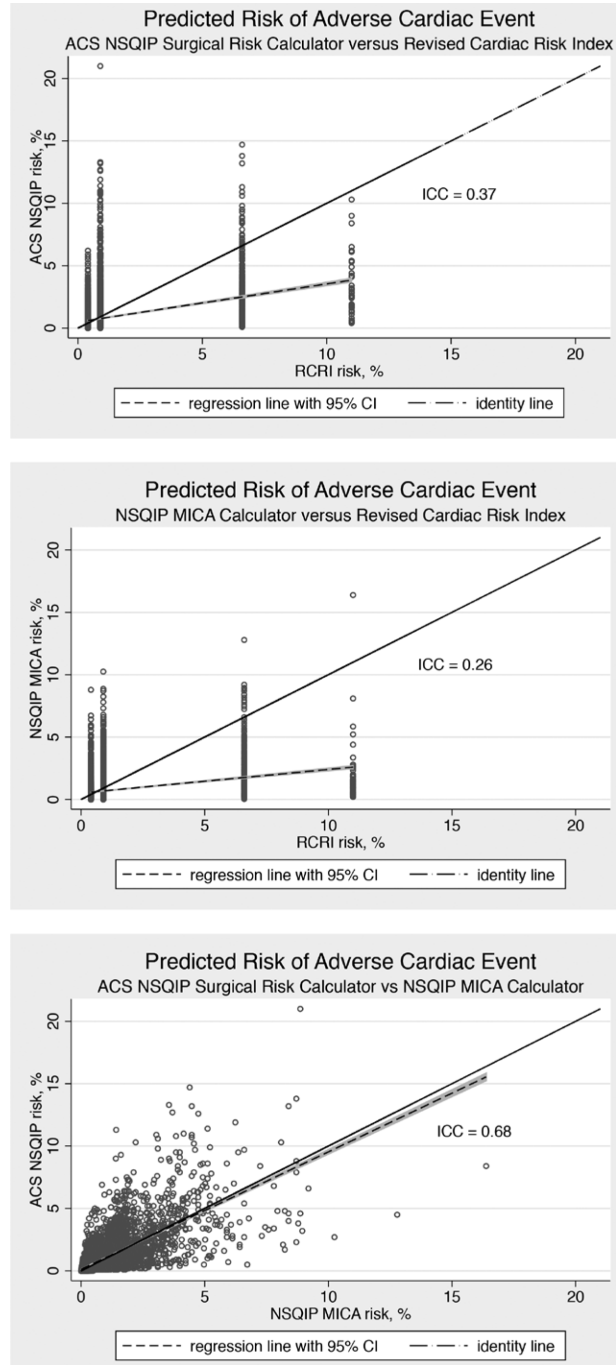


Fig. 3. Comparison of risk of major adverse cardiac event as a function of: (*top*) the American College of Surgeons National Surgical Quality Improvement Program (ACS NSQIP) risk calculator *versus* the Revised Cardiac Risk Index (RCRI); and (*middle*) the National Surgical Quality Improvement Program Myocardial Infarction or Cardiac Arrest (NSQIP MICA) calculator *versus* the RCRI; and (*bottom*) the ACS NSQIP risk calculator *versus* the NSQIP MICA risk calculator. The identity line represents perfect agreement between risk prediction tools. The difference between the identity line and the regression line visually demonstrates the departure from perfect agreement across risk prediction tools. Each point represents a single observation. ICC = intraclass correlation coefficient.

Table 2. Agreement on Low Risk *versus* Elevated Risk: Sensitivity Analyses

	Kappa Statistic				Intraclass Correlation Coefficient		
	Baseline Analysis	Weighted Analysis	Complete Case Analysis	Over-sampling Analysis	Baseline Analysis	Complete Case Analysis	Over-sampling Analysis
ACS NSQIP *vs.* RCRI	0.22	0.25	0.22	0.13	0.37	0.36	0.27
NSQIP MICA *vs.* RCRI	0.18	0.17	0.18	0.12	0.26	0.26	0.21
ACS NSQIP *vs.* NSQIP MICA	0.59	0.53	0.69	0.47	0.68	0.68	0.69
ACS NSQIP *vs.* NSQIP MICA *vs.* RCRI	0.34	NA	0.34	0.26	NA	NA	NA

ACS = American College of Surgeons; MICA = Myocardial Infarction or Cardiac Arrest; NA = not applicable; NSQIP = National Surgical Quality Improvement Program; RCRI = Revised Cardiac Risk Index.

Table 3. Agreement on Which Patients Are Low Risk *versus* Elevated Risk of Major Adverse Cardiac Event as a Function of Which Risk-prediction Tool Is Used to Estimate the Risk of Major Adverse Cardiac Event

	Cases Disagree, (%)			Kappa Analysis		
	Cases Disagree, (%)	95% CI		Kappa	95% CI	
ACS NSQIP *vs.* RCRI	21.6	20.8	22.5	0.22	0.20	0.25
NSQIP MICA *vs.* RCRI	21.5	20.7	22.4	0.18	0.16	0.20
ACS NSQIP *vs.* NSQIP MICA	14.9	14.1	15.6	0.59	0.57	0.61
ACS NSQIP *vs.* NSQIP MICA *vs.* RCRI	29.0	28.1	30.0	0.34	0.33	0.36

ACS = American College of Surgeons; MICA = Myocardial Infarction or Cardiac Arrest; NSQIP = National Surgical Quality Improvement Program; RCRI = Revised Cardiac Risk Index.

Table 4. Agreement on Low-risk *versus* High-risk Patients as a Function of Which Risk-prediction Tool Is Used to Estimate the Risk of Major Adverse Cardiac Event

	RCRI Risk < 1%	RCRI Risk ≥ 1%
ACS NSQIP risk < 1%	6,646 (78.6)	138 (24.7)
ACS NSQIP risk ≥ 1%	1,811 (21.4)	420 (75.3)
NSQIP MICA risk < 1%	6,738 (79.7)	223 (40.0)
NSQIP MICA risk ≥ 1%	1,719 (20.3)	335 (60.0)
	NSQIP MICA Risk < 1%	NSQIP MICA Risk ≥ 1%
ACS NSQIP risk < 1%	6,203 (89.1)	581 (28.3)
ACS NSQIP risk ≥ 1%	758 (10.9)	1,473 (71.7)

Number in parentheses indicates percentage.

ACS = American College of Surgeons; MICA = Myocardial Infarction or Cardiac Arrest; NSQIP = National Surgical Quality Improvement Program; RCRI = Revised Cardiac Risk Index.

database used to create the model; it addresses the generalizability of the risk-prediction tool.[24] Selecting a best-in-class risk calculator for the 2014 American College of Cardiology/American Heart Association perioperative guidelines from one of the currently available models may not be possible for several different reasons. First, the American College of Surgeons National Surgical Quality Improvement Program surgical risk calculator is not in the public domain. Second, the only head-to-head comparison of the National Surgical Quality Improvement Program Myocardial Infarction or Cardiac Arrest risk calculator and the Revised Cardiac Risk Index was conducted with data used to develop the National Surgical Quality Improvement Program Myocardial Infarction or Cardiac Arrest calculator, conferring a performance advantage to the National Surgical Quality Improvement Program Myocardial Infarction or Cardiac Arrest risk calculator. Third, the Revised Cardiac Risk Index, the American College of Surgeons National Surgical Quality Improvement Program risk calculator, and the National Surgical Quality Improvement Program Myocardial Infarction or Cardiac Arrest risk calculators are based on two different definitions of major adverse cardiac event, and are thus, strictly speaking, not comparable. The Revised Cardiac Risk Index defines major adverse cardiac event as myocardial infarction, pulmonary edema, ventricular fibrillation or primary cardiac arrest, or complete heart block.[6] In comparison, the National Surgical Quality Improvement Program definition of major adverse cardiac event, on which the American College of Surgeons National Surgical Quality Improvement Program risk calculator and the National Surgical Quality Improvement Program Myocardial Infarction or Cardiac Arrest risk calculator are based, only includes perioperative myocardial infarction and cardiac arrest.

Our study has several important limitations. *A priori*, one would expect that since the National Surgical Quality Improvement Program risk calculators were developed using National Surgical Quality Improvement Program data, while the Revised Cardiac Risk Index was developed using non–National Surgical Quality Improvement Program data, the National Surgical Quality Improvement Program risk calculators would more accurately predict outcomes than the Revised Cardiac Risk Index in the National Surgical Quality Improvement Program data set. However, it is not likely that this is the main reason that the three risk calculators disagree 30% of the time about which patients were at low-risk of major adverse cardiac event. Instead, it is more likely that these prediction tools disagree because they are fundamentally different models. The Revised Cardiac Risk Index is based on a relatively small sample of patients (4,135 patients) from a single tertiary care center using data that is more than 20 yr old. Further, the Revised Cardiac Risk Index is based on six risk factors, and categorizes procedure-specific risk into high-risk *versus* not high-risk surgery.[6] The American College of Surgeons National Surgical Quality Improvement Program risk calculator, on the other hand, was developed using data from 1.4 million patients in nearly 400 hospitals,[7] has 20 risk factors, classifies procedure-specific risk using a continuous scale, and is updated on a regular basis.

Second, mapping National Surgical Quality Improvement Program data to Revised Cardiac Risk Index risk factors is not perfect, and this could lead to less accurate Revised Cardiac Risk Index predictions, thus contributing to the lack of agreement between the National Surgical Quality Improvement Program risk-prediction tools and the Revised Cardiac Risk Index. However, each of the Revised Cardiac Risk Index risk factors (ischemic heart disease, history of congestive heart failure, history of cerebrovascular disease, insulin therapy for diabetes, and preoperative serum creatinine greater than 2) can be easily mapped to corresponding data elements in National Surgical Quality Improvement Program (*e.g.*, congestive heart failure is a data field in National Surgical Quality Improvement Program). We chose to use American College of Surgeons National Surgical Quality Improvement Program data because it contains the risk factors used in all three risk-prediction tools, and more accurately reflects real world use of these risk-prediction tools than using administrative data in population-based all-payer data (*e.g.*, Nationwide Inpatient Sample). Since the National Surgical Quality Improvement Program is a convenience sample of hospitals, our results are not necessarily generalizable to all U.S. hospitals.

Third, clinicians using risk-prediction tools may have access to more detailed clinical data than the data available in the American College of Surgeons National Surgical Quality Improvement Program database. Many of the risk factors used in these risk-prediction tools are somewhat subjective, and are assigned by surgical clinical reviewers instead of the treating clinician. Fourth, our analysis is incomplete because we cannot compare the results of these risk-prediction tools to clinical judgement alone. Many clinicians may use clinical judgement alone to decide which patients are low risk and can proceed to surgery without further cardiac testing. Nonetheless, our findings have face validity, especially since differences in patient risk estimates across different models are also the root of the reported variability in hospital performance evaluations based on different risk adjustment models.

It is also important to recognize that elevated-risk patients (risk of major adverse cardiac event greater than 1%) will not undergo further testing if: (1) they have moderate or greater (greater than or equal to 4 metabolic equivalents) functional capacity; or (2) if further testing is not expected to impact clinical decision making or perioperative care. In other words, the clinical significance of "elevated risk" is conditional on a patient's functional capacity and other factors.[5] Thus, incorrect estimates of major adverse cardiac event may, in some instances, have no impact on clinical decision making.

The lack of agreement between risk-prediction tools recommended in the American College of Cardiology/American Heart Association guidelines could lead clinicians applying these guidelines to make different clinical decisions depending on which risk-prediction tool is used to stratify patients. Whether these findings are generalizable to other clinical guidelines incorporating more than one risk-prediction tool is an empirical question that requires further investigation. However, it is well recognized that quality metrics based on different risk adjustment models can lead to different conclusions on hospital quality with some hospitals classified as high quality and others as low quality, depending on which risk adjustment model is used.[9,22] The National Quality Forum evaluates the scientific acceptability of quality measures, focusing on the validity of risk adjustment, before endorsing a quality metric.[25] In particular, the National Quality Forum endorses a single best-in-class measure when several quality metrics with the same focus and target population are submitted for evaluation.[26] This avoids the possibility that hospital "quality" will differ for the same outcome and patient population depending on which quality metric is used. Ideally, guideline developers might similarly evaluate the scientific validity of risk-prediction tools (using the available literature) before they are incorporated into clinical algorithms, and attempt to identify a single best-in-class prediction model, whenever possible, to reduce the possibility that a practice guideline might provide clinicians with different recommendations depending on which guideline-recommended prediction tool a clinician selects.

Our findings highlight the need for more accurate prediction modeling for cardiac risk stratification. The venerable Revised Cardiac Risk Index, still the gold standard for cardiac risk stratification, is based on a relatively small patient sample and reflects clinical practice from more than two decades ago. Nonetheless, the Revised Cardiac Risk Index has been extensively validated and exhibits moderate diagnostic accuracy in a recent a meta-analysis based on studies published between 2001 and 2008.[27] The American College of Surgeons National Surgical Quality Improvement Program and the National Surgical Quality Improvement Program Myocardial Infarction or Cardiac Arrest risk calculators, although derived using a very large and contemporary clinical registry, are limited by the narrow definition of major adverse cardiac event (which does not include important cardiac complications such as congestive heart failure) and the possibility that many cardiac complications are not accurately recorded due to incomplete surveillance. Additionally, the National Surgical Quality Improvement Program risk calculators do not include important risk factors such as history of ischemic heart disease or recent myocardial infarction, and have not undergone external validation.

Conclusions

We found wide variability in the predicted risk of major adverse cardiac event for individual patients depending on which risk-prediction tool is used in the 2014 American College of Cardiology/American Heart Association Perioperative Guideline. Including more than one prediction tool in clinical guidelines could lead to differences in decision-making for some patients depending on which risk calculator is used.

Research Support

Supported by funding from the Department of Anesthesiology at the University of Rochester School of Medicine, Rochester, New York.

Competing Interests

The authors declare no competing interests.

Correspondence

Address correspondence to Dr. Glance: University of Rochester Medical Center, 601 Elmwood Avenue, Box 604, Rochester, New York 14642. Laurent_Glance@urmc.rochester.edu. This article may be accessed for personal use at no charge through the Journal Web site, www.anesthesiology.org.

References

1. Weiser TG, Regenbogen SE, Thompson KD, Haynes AB, Lipsitz SR, Berry WR, Gawande AA: An estimation of the global volume of surgery: A modelling strategy based on available data. Lancet 2008; 372:139–44
2. Botto F, Alonso-Coello P, Chan MT, Villar JC, Xavier D, Srinathan S, Guyatt G, Cruz P, Graham M, Wang CY, Berwanger O, Pearse RM, Biccard BM, Abraham V, Malaga G, Hillis GS, Rodseth RN, Cook D, Polanczyk CA, Szczeklik W, Sessler DI, Sheth T, Ackland GL, Leuwer M, Garg AX, Lemanach Y, Pettit S, Heels-Ansdell D, Luratibuse G, Walsh M, Sapsford R, Schünemann HJ, Kurz A, Thomas S, Mrkobrada M, Thabane L, Gerstein H, Paniagua P, Nagele P, Raina P, Yusuf S, Devereaux PJ, Devereaux PJ, Sessler DI, Walsh M, Guyatt G, McQueen MJ, Bhandari M, Cook D, Bosch J, Buckley N, Yusuf S, Chow CK, Hillis GS, Halliwell R, Li S, Lee VW, Mooney J, Polanczyk CA, Furtado MV, Berwanger O, Suzumura E, Santucci E, Leite K, Santo JA, Jardim CA, Cavalcanti AB, Guimaraes HP, Jacka MJ, Graham M, McAlister F, McMurtry S, Townsend D, Pannu N, Bagshaw S, Bessissow A, Bhandari M, Duceppe E, Eikelboom J, Ganame J, Hankinson J, Hill S, Jolly S, Lamy A, Ling E, Magloire P, Pare G, Reddy D, Szalay D, Tittley J, Weitz J, Whitlock R, Darvish-Kazim S, Debeer J, Kavsak P, Kearon C, Mizera R, O'Donnell M, McQueen M, Pinthus J, Ribas S, Simunovic M, Tandon V, Vanhelder T, Winemaker M, Gerstein H, McDonald S, O'Bryne P, Patel A, Paul J, Punthakee Z, Raymer K, Salehian O, Spencer F, Walter S, Worster A, Adili A, Clase C, Cook D, Crowther M, Douketis J, Gangji A, Jackson P, Lim W, Lovrics P, Mazzadi S, Orovan W, Rudkowski J, Soth M, Tiboni M, Acedillo R, Garg A, Hildebrand A, Lam N, Macneil D, Mrkobrada M, Roshanov PS, Srinathan SK, Ramsey C, John PS, Thorlacius L, Siddiqui FS, Grocott HP, McKay A, Lee TW, Amadeo R, Funk D, McDonald H, Zacharias J, Villar JC, Cortés OL, Chaparro MS, Vásquez S, Castañeda A, Ferreira S, Coriat P, Monneret D, Goarin JP, Esteve CI, Royer C, Daas G, Chan MT, Choi GY, Gin T, Lit LC, Xavier D, Sigamani A, Faruqui A, Dhanpal R, Almeida S, Cherian J, Furruqh S, Abraham V, Afzal L, George P, Mala S, Schünemann H, Muti P, Vizza E, Wang CY, Ong GS, Mansor M, Tan AS, Shariffuddin II, Vasanthan V, Hashim NH, Undok AW, Ki U, Lai HY, Ahmad WA, Razack AH, Malaga G, Valderrama-Victoria V, Loza-Herrera JD, De Los Angeles Lazo M, Rotta-Rotta A, Szczeklik W, Sokolowska B, Musial J, Gorka J, Iwaszczuk P, Kozka M, Chwala M, Raczek M, Mrowiecki T, Kaczmarek B, Biccard B, Cassimjee H, Gopalan D, Kisten T, Mugabi A, Naidoo P, Naidoo R, Rodseth R, Skinner D, Torborg A, Paniagua P, Urrutia G, Maestre ML, Santaló M, Gonzalez R, Font A, Martínez C, Pelaez X, De Antonio M, Villamor JM, García JA, Ferré MJ, Popova E, Alonso-Coello P, Garutti I, Cruz P, Fernández C, Palencia M, Díaz S, Del Castillo T, Varela A, de Miguel A, Muñoz M, Piñeiro P, Cusati G, Del Barrio M, Membrillo MJ, Orozco D, Reyes F, Sapsford RJ, Barth J, Scott J, Hall A, Howell S, Lobley M, Woods J, Howard S, Fletcher J, Dewhirst N, Williams C, Rushton A, Welters I, Leuwer M, Pearse R, Ackland G, Khan A, Niebrzegowska E, Benton S, Wragg A, Archbold A, Smith A, McAlees E, Ramballi C, Macdonald N, Januszewska M, Stephens R, Reyes A, Paredes LG, Sultan P, Cain D, Whittle J, Del Arroyo AG, Sessler DI, Kurz A, Sun Z, Finnegan PS, Egan C, Honar H, Shahinyan A, Panjasawatwong K, Fu AY, Wang S, Reineks E, Nagele P, Blood J, Kalin M, Gibson D, Wildes T; Vascular events In noncardiac Surgery patIents cOhort evaluatioN (VISION) Writing Group, on behalf of The Vascular events In noncardiac Surgery patIents cOhort evaluatioN (VISION) Investigators; Appendix 1. The Vascular events In noncardiac Surgery patIents cOhort evaluatioN (VISION) Study Investigators Writing Group; Appendix 2. The Vascular events In noncardiac Surgery patIents cOhort evaluatioN Operations Committee; Vascular events In noncardiac Surgery patIents cOhort evaluatioN VISION Study Investigators: Myocardial injury after noncardiac surgery: A large, international, prospective cohort study establishing diagnostic criteria, characteristics, predictors, and 30-day outcomes. Anesthesiology 2014; 120:564–78
3. Bartels K, Karhausen J, Clambey ET, Grenz A, Eltzschig HK: Perioperative organ injury. Anesthesiology 2013; 119:1474–89
4. Devereaux PJ, Sessler DI: Cardiac complications in patients undergoing major noncardiac surgery. N Engl J Med 2015; 373:2258–69
5. Fleisher LA, Fleischmann KE, Auerbach AD, Barnason SA, Beckman JA, Bozkurt B, Davila-Roman VG, Gerhard-Herman MD, Holly TA, Kane GC, Marine JE, Nelson MT, Spencer CC, Thompson A, Ting HH, Uretsky BF, Wijeysundera DN; American College of Cardiology; American Heart Association: 2014 ACC/AHA guideline on perioperative cardiovascular evaluation and management of patients undergoing noncardiac surgery: A report of the American College of Cardiology/American Heart Association Task Force on practice guidelines. J Am Coll Cardiol 2014; 64:e77–137
6. Lee TH, Marcantonio ER, Mangione CM, Thomas EJ, Polanczyk CA, Cook EF, Sugarbaker DJ, Donaldson MC, Poss R, Ho KK, Ludwig LE, Pedan A, Goldman L: Derivation and prospective validation of a simple index for prediction of cardiac risk of major noncardiac surgery. Circulation 1999; 100:1043–9
7. Bilimoria KY, Liu Y, Paruch JL, Zhou L, Kmiecik TE, Ko CY, Cohen ME: Development and evaluation of the universal ACS NSQIP surgical risk calculator: A decision aid and informed consent tool for patients and surgeons. J Am Coll Surg 2013; 217:833–42.e1–3
8. Gupta PK, Gupta H, Sundaram A, Kaushik M, Fang X, Miller WJ, Esterbrooks DJ, Hunter CB, Pipinos II, Johanning JM, Lynch TG, Forse RA, Mohiuddin SM, Mooss AN: Development and validation of a risk calculator for prediction of cardiac risk after surgery. Circulation 2011; 124:381–7
9. Iezzoni LI: The risks of risk adjustment. JAMA 1997; 278:1600–7

10. User Guide for the 2012 ACS NSQIP Participant Use Data File. Available at: https://www.facs.org/~/media/files/quality%20programs/nsqip/ug12.ashx. Accessed April 13, 2017
11. Fleisher LA, Beckman JA, Brown KA, Calkins H, Chaikof E, Fleischmann KE, Freeman WK, Froehlich JB, Kasper EK, Kersten JR, Riegel B, Robb JF, Smith SC Jr, Jacobs AK, Adams CD, Anderson JL, Antman EM, Buller CE, Creager MA, Ettinger SM, Faxon DP, Fuster V, Halperin JL, Hiratzka LF, Hunt SA, Lytle BW, Nishimura R, Ornato JP, Page RL, Riegel B, Tarkington LG, Yancy CW; ACC/AHA Task Force Members: ACC/AHA 2007 Guidelines on perioperative cardiovascular evaluation and care for noncardiac surgery: Executive summary: A report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (Writing Committee to revise the 2002 guidelines on perioperative cardiovascular evaluation for noncardiac surgery): Developed in collaboration with the American Society of Echocardiography, American Society of Nuclear Cardiology, Heart Rhythm Society, Society of Cardiovascular Anesthesiologists, Society for Cardiovascular Angiography and Interventions, Society for Vascular Medicine and Biology, and Society for Vascular Surgery. Circulation 2007; 116:1971–96
12. Goldzweig CL, Mittman BS, Carter GM, Donyo T, Brook RH, Lee P, Mangione CM: Variations in cataract extraction rates in Medicare prepaid and fee-for-service settings. JAMA 1997; 277:1765–8
13. ACS NSQIP Surgical Risk Calculator, 2018. Available at: https://riskcalculator.facs.org/RiskCalculator/. Accessed March 10, 2018
14. STATA (icc - Intraclass correlation coefficient). College Station, Texas, STATA, 2017
15. Ten Cate DF, Luime JJ, Hazes JM, Jacobs JW, Landewé R: Does the intraclass correlation coefficient always reliably express reliability? Comment on the article by Cheung *et al*. Arthritis Care Res (Hoboken) 2010; 62:1357–8; author reply 1358
16. Cicchetti DV: Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. Psychological Assessment 1994; 6:284
17. Landis JR, Koch GG: The measurement of observer agreement for categorical data. Biometrics 1977; 33:159–74
18. Hosmer DW, Lemeshow S: Applied Logistic Regression. New York, Wiley-Interscience Publication, 2000
19. Cohn SL: Preoperative evaluation for noncardiac surgery. Ann Intern Med 2016; 165:ITC81–96
20. Salisbury AC, Spertus JA: Realizing the potential of clinical risk prediction models: Where are we now and what needs to change to better personalize delivery of care? Circ Cardiovasc Qual Outcomes 2015; 8:332–4
21. Fleisher LA, Beckman JA, Brown KA, Calkins H, Chaikof E, Fleischmann KE, Freeman WK, Froehlich JB, Kasper EK, Kersten JR, Riegel B, Robb JF, Smith SC, Jr., Jacobs AK, Adams CD, Anderson JL, Antman EM, Faxon DP, Fuster V, Halperin JL, Hiratzka LF, Hunt SA, Lytle BW, Nishimura R, Page RL, Riegel B; American Society of Echocardiography; American Society of Nuclear Cardiology; Heart Rhythm Society; Society for Cardiovascular Anesthesiologists; Society for Cardiovascular Angiography and Interventions; Society for Vascular Medicine and Biology: ACC/AHA 2006 guideline update on perioperative cardiovascular evaluation for noncardiac surgery: focused update on perioperative beta-blocker therapy: A report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (Writing Committee to Update the 2002 Guidelines on Perioperative Cardiovascular Evaluation for Noncardiac Surgery) developed in collaboration with the American Society of Echocardiography, American Society of Nuclear Cardiology, Heart Rhythm Society, Society of Cardiovascular Anesthesiologists, Society for Cardiovascular Angiography and Interventions, and Society for Vascular Medicine and Biology. J Am Coll Cardiol 2006; 47: 2343–55
22. Shahian DM, Wolf RE, Iezzoni LI, Kirle L, Normand SL: Variability in the measurement of hospital-wide mortality rates. N Engl J Med 2010; 363:2530–9
23. Cook NR, Ridker PM: Advances in measuring the effect of individual predictors of cardiovascular risk: The role of reclassification measures. Ann Intern Med 2009; 150:795–802
24. Teres D, Lemeshow S: As American as apple pie and APACHE. Acute physiology and chronic health evaluation. Crit Care Med 1998; 26:1297–8
25. Measure Testing and Scientific Acceptability of Measure Properties, 2011. Available at: https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=3&cad=rja&uact=8&ved=0ahUKEwiG5Kfp9qLXAhXMBsAKHZ2IBqUQFgg4MAI&url=https%3A%2F%2Fwww.qualityforum.org%2FMeasuring_Performance%2FSubmitting_Standards%2FMeasure_Testing_and_Scientific_Acceptability.aspx&usg=AOvVaw03JmEiMkWiujOPULNZ5N6M. Accessed November 4, 2017
26. National Quality Forum Guidance on Competing Measures and Selection of the Best Measure, 2011. Available at: https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=5&ved=0ahUKEwj2x6DT7qLXAhXFLyYKHbaUByUQFgg9MAQ&url=https%3A%2F%2Fwww.qualityforum.org%2FWorkArea%2Flinkit.aspx%3FLinkIdentifier%3Did%26ItemID%3D69378&usg=AOvVaw2vJYoDAcqGDfLQtcuxYj-U. Accessed November 4, 2017
27. Ford MK, Beattie WS, Wijeysundera DN: Systematic review: Prediction of perioperative cardiac complications and mortality by the revised cardiac risk index. Ann Intern Med 2010; 152:26–35

Appendix: Definition of Major Adverse Cardiac Event (MACE)

Risk-prediction tool	Definition of MACE
Revised Cardiac Risk Index	Myocardial infarction Cardiac arrest Pulmonary edema (based on a "reading of chest radiograph consistent with this complication in a plausible clinical setting.") Ventricular fibrillation Complete heart block
ACS NSQIP Risk Calculator	Myocardial infarction: (1) "documentation of ECG changes indicative of acute MI (ST elevation greater than 1 mm in two contiguous leads, new left bundle branch block, or new Q waves in two or more contiguous leads)"; (2) "new elevation of troponin greater than three times the upper level of the reference range in the setting of suspected myocardial ischemia"; or (3) "physician diagnosis of myocardial infarction." (User Guide for the 2012 ACS NSQIP Participant Use Data File) Cardiac arrest: "the absence of cardiac rhythm or presence of chaotic cardiac rhythm, intraoperatively or within 30 days following surgery, which results in a cardiac arrest requiring the initiation of CPR, which includes chest compressions. Patients are included who are in a pulseless VT or Vfib in which defibrillation is performed and PEA arrests requiring chest compressions. Patients with automatic implantable cardioverter defibrillator that fire but the patient has no loss of consciousness should be excluded." (User Guide for the 2012 American College of Surgeons National Surgical Quality Improvement Program Participant Use Data File)
NSQIP MICA Risk Calculator	Myocardial infarction (same definition as ACS NSQIP Risk Calculator) Cardiac arrest (same definition as ACS NSQIP Risk Calculator)

ACS NSQIP = American College of Surgeons National Surgical Quality Improvement Program; CPR = cardiopulmonary resuscitation; ECG = electrocardiogram; MI = myocardial infarction; NSQIP MICA = National Surgical Quality Improvement Program Myocardial Infarction or Cardiac Arrest; PEA = pulseless electrical activity; Vfib = ventricular fibrillation; VT = ventricular tachycardia.