Validation of a Risk Stratification Index and Risk Quantification Index for Predicting Patient Outcomes

In-hospital Mortality, 30-day Mortality, 1-year Mortality, and Length-of-stay

Matthew J. G. Sigakis, M.D.,* Edward A. Bittner, M.D., Ph.D.,† Jonathan P. Wanderer, M.D., M. Phil.‡

ABSTRACT

Background: External validation of published risk stratification models is essential to determine their generalizability. This study evaluates the performance of the Risk Stratification Indices (RSIs) and 30-day mortality Risk Quantification Index (RQI).

Methods: 108,423 adult hospital admissions with anesthetics were identified (2006–2011). RSIs for mortality and length-of-stay endpoints were calculated using published methodology. 91,128 adult, noncardiac inpatient surgeries were identified with administrative data required for RQI calculation.

Results: RSI in-hospital mortality and RQI 30-day mortality Brier scores were 0.308 and 0.017, respectively. RSI discrimination, by area under the receiver operating curves, was excellent at 0.966 (95% CI, 0.963–0.970) for in-hospital mortality, 0.903 (0.896–0.909) for 30-day mortality, 0.866 (0.861–0.870) for 1-yr mortality, and 0.884 (0.882–0.886) for length-of-stay. RSI calibration, however, was poor overall (17% predicted in-hospital mortality *vs.* 1.5% observed after inclusion of the regression constant) as demonstrated by calibration plots. Removal of self-fulfilling diagnosis and procedure codes (20,001 of 108,423; 20%) yielded similar results. RQIs were calculated for only 62,640 of 91,128 patients (68.7%) due to unmatched procedure codes. Patients with unmatched codes were younger, had higher American Society of Anesthesiologists physical status and 30-day mortality. The area under the receiver operating curve for 30-day mortality RQI was 0.888 (0.879–0.897). The model also demonstrated good calibration. Performance of a restricted index, Procedure Severity Score + American Society of Anesthesiologists physical status, performed as well as the original RQI model (age + American Society of Anesthesiologists + Procedure Severity Score).

Conclusion: Although the RSIs demonstrated excellent discrimination, poor calibration limits their generalizability. The 30-day mortality RQI performed well with age providing a limited contribution.

What We Already Know about This Topic

- Risk Stratification Indices and Risk Quantification Indices were developed to predict clinical endpoints using administrative patient data
- External validation of risk stratification models determine their generalizability
- Validation should characterize overall performance and assess discrimination (the probability measured risk is higher for a case than it is for a noncase) and calibration (how well-predicted probabilities align with observed outcomes)

What This Article Tells Us That Is New

- Patient data from the Massachusetts General Hospital were used to show that the Risk Stratification Indices had excellent discrimination and poor calibration but the 30-day mortality Risk Quantification Indices performed well

ADMINISTRATIVE patient data are increasingly used to evaluate patterns and outcomes of disease. Several indices have been developed to predict mortality and other endpoints, including the Charlson Comorbidity Index,[1] variations on the Charlson Comorbidity Index,[2,3] the Elixhauser method,[4] and the Procedural Index for Mortality Risk.[5] The Risk Stratification Indices[6] (RSIs) were developed using International Classification of Disease, Ninth

* Chief Resident, † Program Director, Critical Care Fellowship, Assistant Professor, Department of Anesthesia, Critical Care, and Pain Medicine, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts. ‡ Instructor, Department of Anesthesia, Vanderbilt University, Nashville, Tennessee.

Received from the Department of Anesthesia, Critical Care, and Pain Medicine, Massachusetts General Hospital, Boston, Massachusetts, and Department of Anesthesia, Vanderbilt University, Nashville, Tennessee. Submitted for publication July 13, 2012. Accepted for publication April 17, 2013. Support was provided solely from institutional and/or departmental sources. This work in part has been presented at the Annual Meeting of the American Society of Anesthesiologists, Washington, DC, October 13, 2012. The authors declare no competing interests.

Address correspondence to Dr. Sigakis: Department of Anesthesia, Critical Care, and Pain Medicine, Massachusetts General Hospital, Gray-Bigelow 444, 55 Fruit Street, Boston, Massachusetts 02114. msigakis@partners.org. This article may be accessed for personal use at no charge through the Journal Web site, www.anesthesiology.org.

 Anesthesiology 2013; 119:525-40

◇ This article is featured in "This Month in Anesthesiology." Please see this issue of ANESTHESIOLOGY, page 3A.

Revision, Clinical Modification (ICD-9-CM) diagnosis and procedure codes for hospital inpatients who were more than 65 yr of age, which were obtained from the Medicare Provider Analysis and Review (MEDPAR) database for the period of 2001–2006. The RSIs have received much attention for their potential application as a tool for risk-adjusting important healthcare outcomes.[7] However, they have not yet been externally validated.

In the RSI models, a level of risk was derived for diagnosis and procedure codes associated with a hospital stay, using logistic regression modeling for in-hospital mortality and Cox proportional hazards modeling for postdischarge death and length-of-stay (LOS). Summation of the covariate coefficients associated with a patient's procedure and diagnostic codes generates an RSI. The RSIs demonstrated excellent discrimination (c statistic) as predictive indices for mortality and LOS endpoints when applied to a large MEDPAR validation set and the authors' institutional data. Additional statistical analysis to determine overall model performance and calibration was recommended.[8] Calibration of a model refers to the degree to which predicted and actual outcomes agree.[9] A follow-up analysis assessed calibration by comparing actual mortality rates with RSI value groups.[10] A comparison against predicted mortality was not made.

The Risk Quantification Indices[11] (RQIs) are another risk index system developed to predict 30-day mortality and morbidity using a small number of administrative data points. The indices were developed using adult noncardiac surgical patient data from the National Surgical Quality Improvement Project database[12] for the period of 2005–2008. Using Current Procedural Terminology (CPT) codes, the authors derived a procedural severity score (PSS) that measured procedural risk. These scores were combined with patient age and American Society of Anesthesiologists (ASA) physical status to create a predictive index of 30-day mortality and major morbidity. The goal of this study was to externally evaluate the performance of the RSIs and 30-day mortality RQI using patient data at our institution.

Materials and Methods

Patient Population

This study was approved by the Partners Institutional Review Board, Boston, Massachusetts (2011P000253). For validation of the RSIs, the authors identified adult (18 yr of age or older) inpatient admissions that included anesthetics for the period 2006–2011 using the Massachusetts General Hospital (MGH) anesthesia information management system. For the 30-day mortality RQI validation, we identified all adult noncardiac inpatient surgical cases by excluding cardiac cases and nonsurgical procedures that required anesthetics, specifically cardiology, electroconvulsive therapy, and labor epidurals.

§ Cleveland Clinic: Outcomes Research, Risk Stratification Index. Available at: http://my.clevelandclinic.org/anesthesia/outcomes/risk-stratification-index.aspx. Accessed November 23, 2012.

Table 1. RSI Dataset Characterization

	RSI Data Set
Total cases in data set	108,423
Age, yr	54.5 ± 19.9
Female, %	53.5
White/black/other, %	82.3/4.3/13.4
Dx count	6.9 ± 2.9
Px count	3.0 ± 2.4
LOS, d	6.6 ± 9.6
In-hospital mortality, %	1.1
30-d mortality, %	2.0
1-yr mortality, %	6.5

Data are presented as mean ± SD unless noted otherwise.
Dx = ICD-9-CM Diagnosis code; ICD-9-CM = International Classification of Disease, Ninth Revision, Clinical Modification; LOS = length of stay; Px = ICD-9-CM Procedure code; RSI = Risk Stratification Index.

Data Collection

After identifying the patient population using MGH anesthesia information management system data, we obtained the ASA physical status and primary surgical CPT code from our anesthesia billing system (2007–2011). Diagnostic and procedure ICD-9-CM codes and LOS data were obtained from the MGH billing system. Mortality endpoints were derived from the Partners HealthCare System Research Patient Data Repository (RPDR). The RPDR is a secure centralized administrative data warehouse that contains patient encounter data from multiple hospital information systems in the Partners Health System.[13,14] The RPDR links patients to the National Death Index, a central computerized index of death record information maintained by the National Center for Health Statistics division of the Centers for Disease Control.[15]

RSI Validation Methodology

For validation of the RSIs, a level of risk was assigned for each patient in our sample population by summation of the covariate coefficients associated with each diagnosis and procedure ICD-9-CM code. The methodology for RSI calculation is available at the authors' Web site.§ Briefly, patient data are organized into an "input file," with each row representing a single patient stay that contains all associated ICD-9-CM diagnosis and procedure codes. Using SPSS (version 17.0; IBM, Armonk, NY), the published SPSS macro was executed with the MGH input file to assign β coefficients to each diagnosis and procedure code. β Coefficients were summed to calculate an RSI value.

Overall model performance was determined using the Brier score. To assess discrimination, the aggregate performance of each RSI on outcomes of interest was quantified by calculating the area under the receiver operating characteristic curve (AUROC). Consistent with the original RSI study, LOS was assessed by determining whether the index risk was above or below the median LOS for the primary ICD-9-CM procedure code. Calibration was assessed by comparing

Table 2. Most Common ICD-9-CM Procedure and Diagnosis Codes

	Count	%	Definition
Procedure code			
P741	3,528	3.3	Low cervical cesarean section
P8154	3,324	3.1	Total knee replacement
P7569	3,193	2.9	Repair of other current obstetric laceration
P8151	3,039	2.8	Total hip replacement
P605	2,136	2.0	Radical prostatectomy
P0309	1,521	1.4	Other exploration and decompression of spinal canal
P3521	1,517	1.4	Replacement of aortic valve with tissue graft
P7936	1,509	1.4	Open reduction of fracture with internal fixation, tibia, and fibula
P3812	1,465	1.4	Endarterectomy, other vessels of head and neck
P064	1,423	1.3	Complete thyroidectomy
Diagnosis code			
D71536	2,952	2.7	Osteoarthrosis, localized, not specified whether primary or secondary, lower leg
D71535	2,340	2.2	Osteoarthrosis, localized, not specified whether primary or secondary, pelvic region, and thigh
D185	2,214	2.0	Other lung diseases
D66411	1,677	1.5	Second-degree perineal laceration, delivered, with or without mention of antepartum condition
D41401	1,604	1.5	Coronary atherosclerosis of native coronary artery
D42731	1,365	1.3	Atrial fibrillation
D43310	1,286	1.2	Occlusion and stenosis of carotid artery without mention of cerebral infarction
D27801	1,277	1.2	Morbid obesity
D65421	1,188	1.1	Previous cesarean delivery, delivered, with or without mention of antepartum condition
D4241	1,104	1.0	Other endocardial disease

ICD-9-CM = International Classification of Disease, Ninth Revision, Clinical Modification.

actual with predicted endpoints[16] for in-hospital mortality, 30-day mortality, and 1-yr mortality in the RPDR data set. Predicted outcomes were calculated by the inverse logit function for mortality endpoints and LOS ($1/[1+e^{-RSI}]$). For time-dependent endpoints (30-day mortality, 1-yr mortality, and LOS), calibration curves were generated using cases where endpoints were known, eliminating the concern for censored data. The authors defined "self-fulfilling" ICD-9-CM codes as "conditions that required immediate medical intervention" and "procedures that are typically performed during emergency resuscitation," and then applied this definition to the RSI in-hospital mortality model (appendix 1). ICD-9-CM codes that are self-fulfilling (*i.e.*, cardiac arrest) with respect to outcome were removed from the MGH data set and RSI model performance reanalyzed.

RQI Validation Methodology

To validate the RQI for 30-day mortality, CPT codes corresponding to patients' primary procedure were assigned weights (*i.e.*, PSS) and combined with ASA physical status and age to calculate the RQI. Technical issues precluded the use of the published R module, available at the RQI Web site.‖ Published PSS covariates together with model parameters provided by the original authors (appendix 2) were used to compute the 30-day mortality RQI using SPSS (appendix 3). Results were compared against actual patient endpoints in the RPDR data set. Overall model performance was assessed with the Brier score. Discrimination was quantified by AUROC calculation. Calibration was assessed by comparing actual with predicted endpoints[16] for 30-day mortality in the RPDR data set. RQI model performance (age, ASA physical status, and PSS) was compared with limited versions (age + ASA, age + PSS, ASA + PSS) to assess the degree to which each variable contributed to the RQI performance.

Statistical Analysis

In addition to the validation methodology above, statistical comparisons were made to characterize the patient population that did not generate RQI values due to unmatched CPT codes. The chi-square test was used for categorical variables (mortality, ASA physical status >2) and a two-tailed independent *t* test was used to compare continuous variables (age). All statistical comparisons were performed using SPSS. Calibration plots and Brier scores were generated using R (version 2.15.1, rms package; R Core Team, R Foundation for Statistical Computing, Vienna, Austria).#

‖ Cleveland Clinic: Outcomes Research, Risk Quantification Index. Available at: http://my.clevelandclinic.org/anesthesia/outcomes/risk-quantification-index.aspx. Accessed November 23, 2012.

Vanderbilt University, Department of Biostatistics: Statistical Computing. Available at: http://biostat.mc.vanderbilt.edu/wiki/Main/StatComp. Accessed November 23, 2012.

Table 3. RSI and RQI Brier Scores and AUROCs

	Outcome	MGH Brier Score	MGH AUROC	Sessler *et al.* AUROC
RSI	In-hospital mortality	0.308	0.966 (0.963–0.970)	0.977 (0.977–0.980)
	Adjusted in-hospital mortality	0.073	0.966 (0.963–0.970)	
	30-d mortality	0.199	0.903 (0.896–0.909)	0.854 (0.834–0.875)
	1-yr mortality	0.195	0.866 (0.861–0.870)	0.832 (0.825–0.839)
	LOS	0.147	0.884 (0.882–0.886)	0.886 (0.883–0.888)
	In-hospital mortality w/o SFC	0.253	0.973 (0.968–0.982)	
	Adjusted in-hospital mortality w/o SFC	0.036	0.973 (0.968–0.982)	
				Dalton *et al.* AUROC
RQI	30-d mortality	0.017	0.888 (0.879–0.897)	0.915 (0.906–0.924)

AUROC = area under the receiver operating characteristic; ICD-9-CM = International Classification of Disease, Ninth Revision, Clinical Modification; LOS = length-of-stay; MGH = Massachusetts General Hospital; RQI = Risk Quantification Index; RSI = Risk Stratification Index; SFC = self-fulfilling ICD-9-CM codes; w/o = without.

The specific code used to generate RSI and RQI calibration plots can be found in appendix 4.

Results

RSI

A total of 108,423 adult anesthetic records, each corresponding to an individual inpatient admission, were identified for validating the RSIs. Overall, there were 3,811 unique principal diagnosis codes and 1,873 unique principal procedure codes. Characteristics of the data set are illustrated in tables 1 and 2. Overall, the RSIs demonstrated excellent discrimination, but poor calibration. The results for each endpoint's Brier score and AUROC are summarized in table 3 and figure 1. The AUROC for in-hospital mortality, 30-day

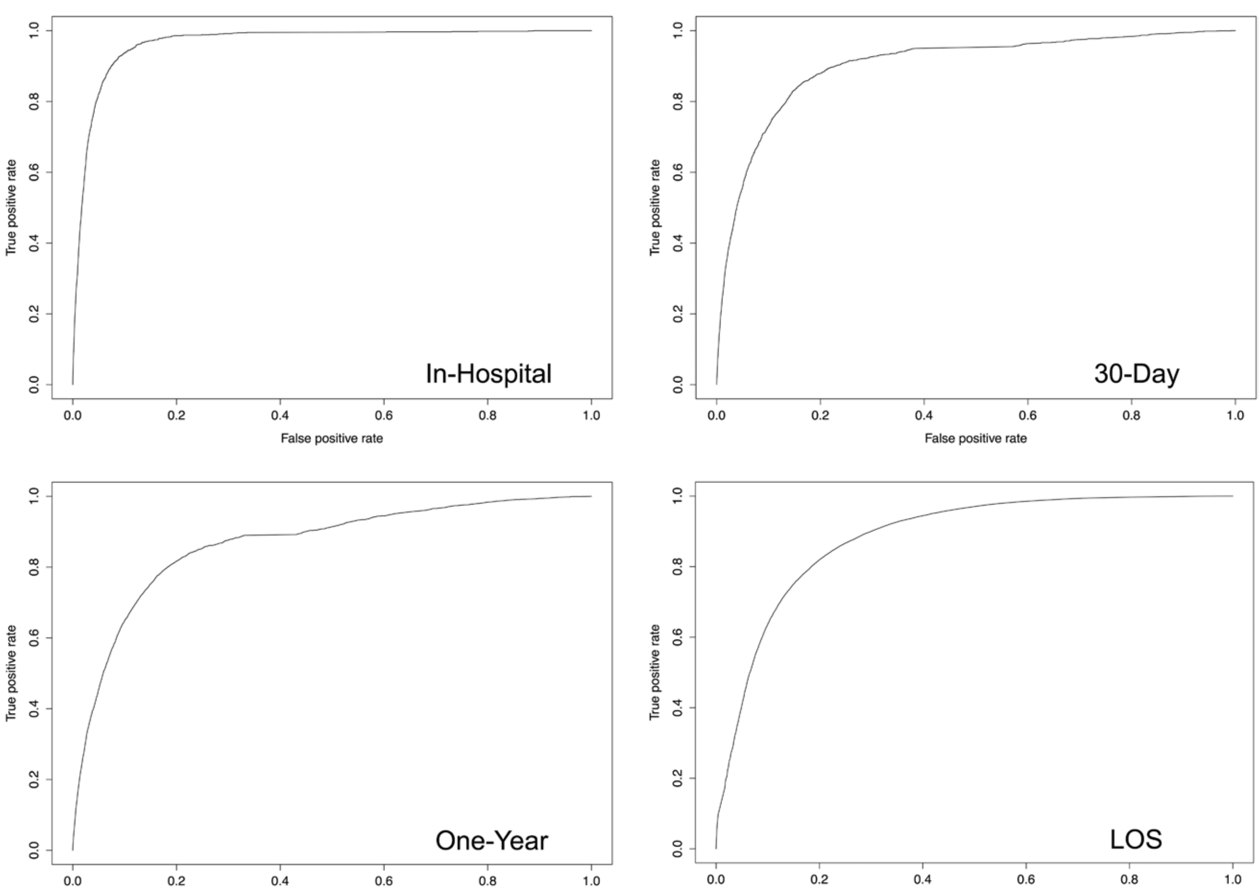

Fig. 1. External validation of Risk Stratification Index discrimination: area under the receiver operating characteristic curves for Risk Stratification Index endpoints using Massachusetts General Hospital patient data. In-Hospital = in-hospital mortality; LOS = length-of-stay; One-Year = 1-year mortality; 30-Day = 30-day mortality.

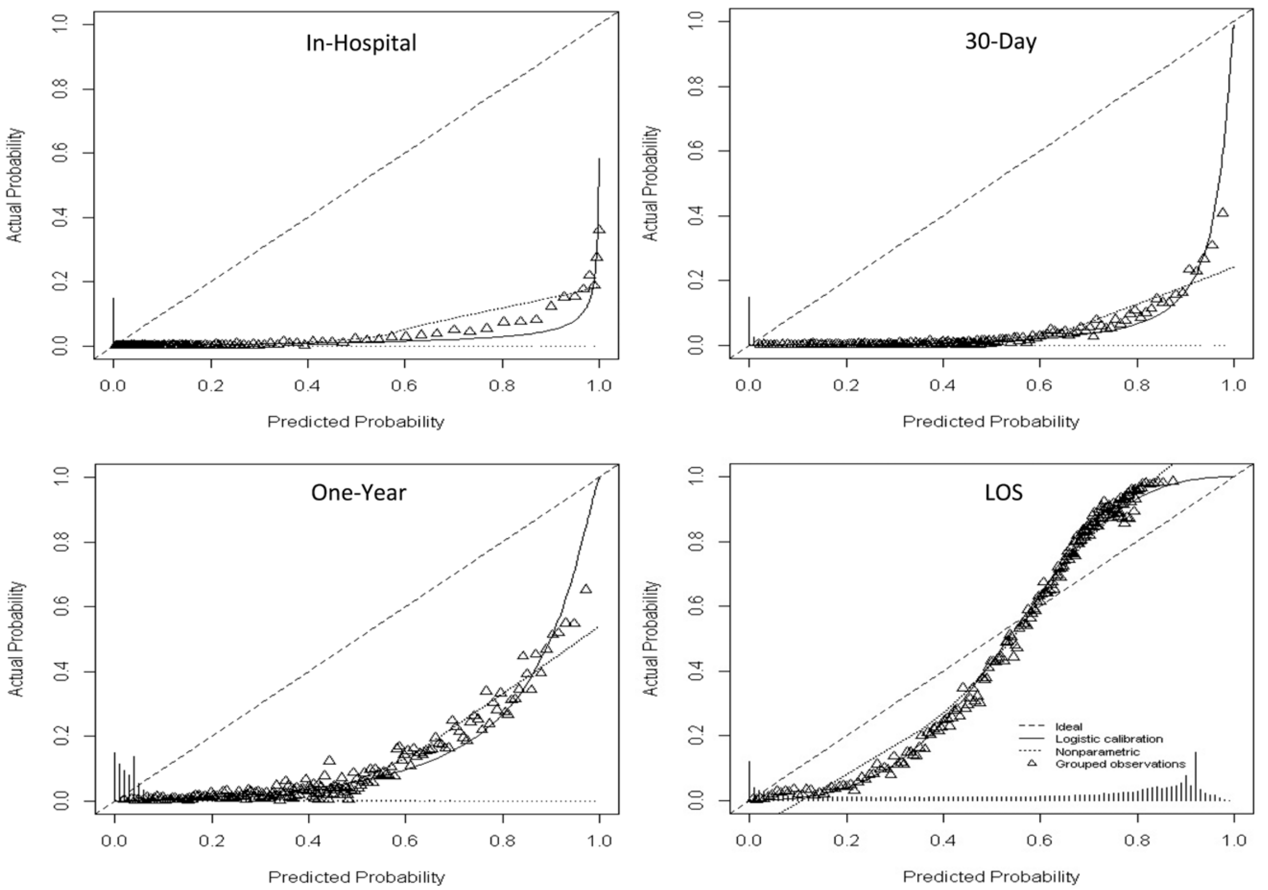

Fig. 2. External validation of Risk Stratification Index calibration: calibration curves for Risk Stratification Index endpoints using Massachusetts General Hospital patient data. In-Hospital = in-hospital mortality; LOS = length-of-stay; One-Year = 1-year mortality; 30-Day = 30-day mortality.

mortality, 1-yr mortality, and LOS at our institution were 0.966 (95% CI, 0.963–0.970), 0.903 (0.896–0.909), 0.866 (0.861–0.870), and 0.884 (0.882–0.886), respectively, compared with the original findings of 0.977 (0.977–0.980), 0.854 (0.834–0.875), 0.832 (0.825–0.839), and 0.886 (0.883–0.888), respectively.[6]

Calibration "in-the-large" for RSI in-hospital mortality illustrated a discrepancy between actual (1.5%) and predicted (51.7%) in-hospital mortality. The authors identified a regression constant (–2.198) in the published RSI "all-covariates.xls" file§ that was not used in the published SPSS implementation macro, "RSI Calculation for Web Use Rev

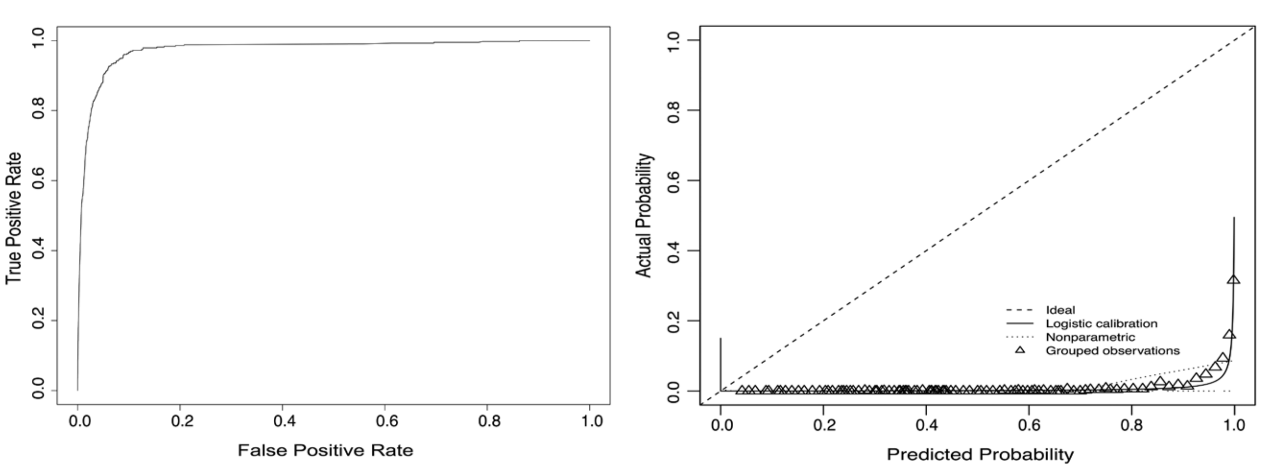

Fig. 3. Risk Stratification Index in-hospital mortality discrimination and calibration after removal of "self-fulfilling" International Classification of Diseases, Ninth Revision, Clinical Modification codes from Massachusetts General Hospital patient data: receiver operating characteristic curve and calibration plot.

Table 4. RQI Data Set Characterization

	RQI Data Set(Input set)	RQI Data Set (Calculated)	RQI Data Set(Missing)
Total cases in data set	91,128	62,640	28,488
Age, yr	54.5 ± 20.0	56.2 ± 19.1	50.8 ± 21.6*
ASA I/II/III/IV/V, %	8.0/50.9/34.8/6.1/0.2	8.2/55.5/31.8/4.3/0.3	7.4/40.8/41.4/10.3/0.2
30-d mortality, %	2.8	1.9	4.7*
High ASA, %	43.1	38.0	54.3*

*$P < 0.001$ compared with RQI data set (calculated) using two-tailed independent *t* test for continuous variable (age) and chi-square test for categorical variables (mortality, high ASA). Data are presented as mean ± SD unless noted otherwise.

ASA = American Society of Anesthesiologists physical status; Calculated = cases with CPT codes where a PSS could be calculated based on published RQI methodology; CPT = current procedural terminology; High ASA = ASA > II; Input set = cases with the data required for calculation of the 30-d mortality RQI; Missing = cases with unmatched CPT codes preventing calculation of 30-day mortality RQI; PSS = Procedure Severity Score; RQI = Risk Quantification Index.

2.sps,"§ and used this to calculate an *adjusted* RSI in-hospital mortality. Incorporation of this constant improved the calibration (predicted in-hospital mortality of 17.5%) and Brier score (table 3), although calibration remained poor. Calibration plots also demonstrated poor calibration for 30-day mortality, 1-yr mortality, and LOS (fig. 2). Assessment of a sample results data file published on the RSI Web site, "sample data rev2.sav,"§ was similar (49% predicted in-hospital mortality, 18% predicted in-hospital mortality after including the regression constant).

Records with any ICD-9-CM codes that were identified as self-fulfilling with respect to outcome (appendix 1) were removed from the MGH data set (20,001 of 108,423, 20%) and RSI model performance reanalyzed; discrimination slightly improved whereas calibration remained poor (fig. 3).

RQI

The authors identified 91,128 anesthetic records for noncardiac surgical cases with the data required for calculation of the 30-day mortality RQI. There were 1,694 unique primary surgical CPT codes. Characteristics of the data set are shown in table 4. Of these, RQI calculations could not be performed for 28,488 cases (31.3%) due to unmatched CPT codes. Of the 197 unique unmatched CPT codes, the 10 most common are illustrated in table 5. Compared with the matched CPT data set, patients with unmatched CPT codes were younger (50.8 ± 21.6 *vs.* 56.2 ± 19.1 yr old; $P < 0.001$),

Table 5. Most Common Unmatched CPT Codes

Primary Surgical CPT	Count	%	CPT Description
59514AA	1,903	6.7	Cesarean delivery only
59514AC	851	3.0	Cesarean delivery only
61624AA	703	2.5	Transcatheter permanent occlusion or embolization (*e.g.*, for tumor destruction, to achieve hemostasis, to occlude a vascular malformation), percutaneous, any method; central nervous system (intracranial, spinal cord)
43235AA	616	2.2	Upper gastrointestinal endoscopy including esophagus, stomach, and either the duodenum and/or jejunum as appropriate; diagnostic, with or without collection of specimen(s) by brushing or washing (separate procedure)
43246AA	599	2.1	Upper gastrointestinal endoscopy including esophagus, stomach, and either the duodenum and/or jejunum as appropriate; with directed placement of percutaneous gastrostomy tube
31600AA	525	1.8	Tracheostomy, planned (separate procedure)
52332AA	513	1.8	Cystourethroscopy, with insertion of indwelling ureteral stent (*e.g.*, Gibbons or double-J type)
45380AA	497	1.7	Colonoscopy, flexible, proximal to splenic flexure; with biopsy, single or multiple
45378AA	476	1.7	Colonoscopy, flexible, proximal to splenic flexure; diagnostic, with or without collection of specimen(s) by brushing or washing, with or without colon decompression (separate procedure)
31622AA	469	1.7	Bronchoscopy, rigid or flexible, including fluoroscopic guidance, when performed; diagnostic, with cell washing, when performed (separate procedure)

CPT = Current Procedural Terminology.

were more likely to have an ASA physical status greater than 2 (54.3 *vs.* 38.0%; $P < 0.001$) and had greater 30-day mortality (4.7 *vs.* 1.9%; $P < 0.001$).

For the 62,640 cases with matched CPT codes, Brier score and AUROC are shown in table 3 and figure 4. AUROC for 30-day mortality at our institution was 0.888 (0.879–0.897). This performance was similar to the originally reported AUROC of 0.915 (0.906–0.924).[11] The 30-day mortality RQI demonstrated good calibration (fig. 4).

The discrimination of the age + ASA, age + PSS, and ASA + PSS models are shown in table 6. Calibration of these component models was not assessed. Of the three model elements, age provided the smallest contribution to the resulting discrimination.

Discussion

Our results indicate important limitations to the generalizability of RSIs and 30-day mortality RQI. To the authors' knowledge, this is the first comprehensive external evaluation of these indices to be published.

The RSIs use ICD-9-CM codes for diseases and procedures associated with hospitalization. Using up to 10 diagnostic and 10 procedure codes to assign a level of risk to a patient stay, the RSIs capture the underlying clinical condition in a manner that is intended for retrospective risk-adjusted quality-of-care comparisons. By contrast, the RQIs use information available before a hospitalization to predict expected outcomes. Compared with other indices for outcome adjustment, the RQI uses fewer data points: a primary surgical CPT code, ASA physical status, and age.

Both RSI and RQI indices use an aggregation scheme to account for low-frequency codes and annual revisions in code definitions. In the RSI models, hierarchical selection processes were used on the MEDPAR data set to select a set of codes based on average annual incidence to ensure consistency of codes across years. Less frequent ICD-9-CM codes were truncated and reassessed to ensure an average annual occurrence of more than 5,000 for four- and five-character ICD-9-CM codes and more than 1,000 for three-character ICD-9-CM codes. ICD-10, which consists of over 100,000 α-numeric diagnostic and procedure codes, is scheduled for roll out in October 2014.[17] A similar derivation scheme could be applied to generate RSI models compatible with ICD-10. The RQI aggregation scheme was developed on a reserved subset of the National Surgical Quality Improvement Project database.[12] Frequently used CPT codes were represented as separate cohorts. Less common procedures were aggregated according to one of 244 categories as described by the Clinical Classifications Software for Services and Procedures (U.S. Department of Health and Human Services Agency for Healthcare Research and Quality, Rockville, MD).[18] If the number of cases within a Clinical Classifications Software group was low, these groups were further aggregated into an "all-purpose other" group. The resulting scheme defines PSS scores for a subset of CPT codes.

Evaluation of a predictive model should characterize overall performance, and assess discrimination *and* calibration.[8,16,19] The Brier score is a statistical measure used to assess overall predictive performance by computing the squared difference between predicted and actual outcomes.[16,20] Smaller differences between predicted and actual data points reflect a better overall "goodness-of-fit" of a model. The Brier score ranges from 0 to 1. A model with perfect prediction has a score of 0, whereas a noninformative model that provides predictions no better than chance with an outcome incidence of 50% has a score of 0.25. Unlike the AUROC, interpretation of the Brier score depends on the incidence of the outcome. The Adjusted RSI In-Hospital Mortality model without self-fulfilling codes had the best overall Brier score relative to other RSIs, but was still considerably higher than the RQI 30-day Mortality Brier score.

Calibration refers to how well a model's predicted probabilities align with the observed outcomes. Discrimination, however, represents the probability that the measured risk is higher for a case *versus* noncase or how well a model can rank order cases.[9] It is not surprising, for instance, that discrimination for RSI In-Hospital Mortality remained strong after removal of self-fulfilling ICD-9-CM codes. A well-ordered list will remain in order regardless of how cases may be removed. But discrimination does not characterize actual predicted probabilities, which are fundamental to clinical risk-prediction models. Calibration is especially important for prognostic models in which the clinical question is the chance of a future outcome, given current risk factors.[16,19] RSI In-Hospital Mortality calibration remained poor after removal of self-fulfilling ICD-9-CM codes. The improvement in Brier score likely reflects the modest improvement in discrimination.

Although the RSIs demonstrate excellent discrimination, the current models' poor calibration limits their use as a tool to compare risk-adjusted outcomes among entities. Of note, there is a significant mean age difference between the MEDPAR data set (mean age 74) and MGH data set (mean age 55). This age difference may influence calibration, as greater risk is associated with increasing age. Although the original RSI study includes an internal validation of Cleveland Clinic data with mean age of 56.6 yr,[6] there are no calibration data published for comparison. The physiologic changes of aging result in increased incidence of comorbidities such as high blood pressure** and diabetes.†† For example, the impact of diabetes on recovery after hip fracture is moderated by age.[21] Thus, the risk associated with a diagnosis or procedure

** American Heart Association: Statistical Fact Sheet (2012). Available at: http://www.heart.org/HEARTORG/General/Populations_UCM_319119_Article.jsp. Accessed June 10, 2013.

†† American Diabetes Association: Diabetes Statistics (2011). Available at: http://www.diabetes.org/diabetes-basics/diabetes-statistics/. Accessed February 2, 2013.

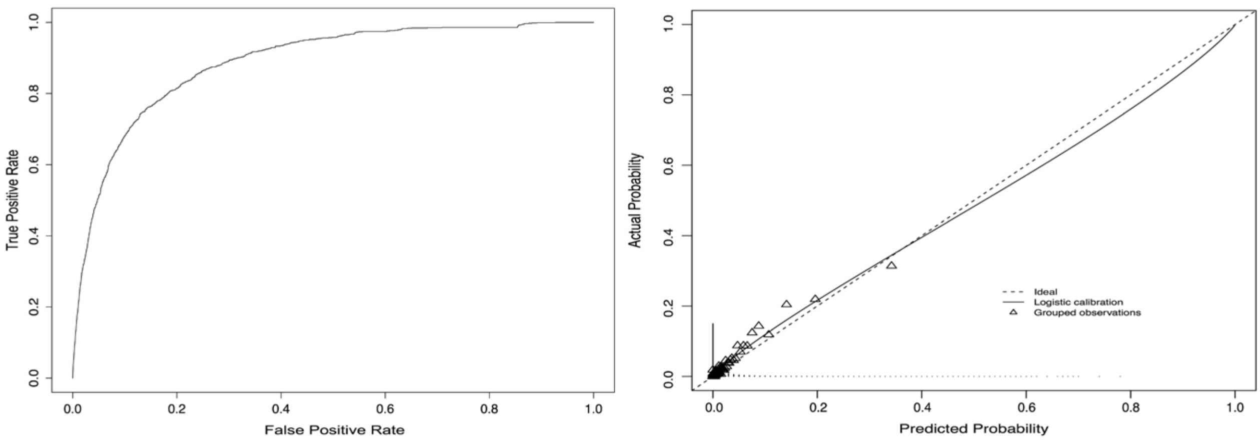

Fig. 4. External validation of Risk Quantification Index 30-day mortality discrimination and calibration: Area under the receiver operating characteristic curve and calibration curve for Risk Quantification Index 30-day mortality endpoint using Massachusetts General Hospital patient data.

code may be different when comparing patients based on age alone.

It is interesting to note the impact of age in another well-established predictive index, the Revised Cardiac Risk Index. In the Revised Cardiac Risk Index, discrimination was evaluated using the AUROC whereas calibration was assessed by comparing the predicted and actual major cardiac complication rates by risk class.[22] The mean age of patients within the Revised Cardiac Risk Index study was 66 yr, with age more than 70 yr correlating with a relative risk of 1.9 (1.1–3.2) for cardiac complications.[22] Although increased age was associated with higher morbidity in the original Cardiac Risk Index,[23] the six independent risk factors identified by logistic regression analysis for the Revised Cardiac Risk Index did not include age.[22,24]

For the RQI calculation, we observed a relatively high rate of unmatched CPT codes in our data set of inpatient anesthetics, which is likely a product of the aggregation scheme and the relatively broad set of anesthetic cases included in our data set. Alternative CPT aggregation schemes have been proposed.[25] The data set used to derive the RQI model was obtained from the National Surgical Quality Improvement Project and may not be representative of procedures or CPT coding practices at our institution. Our data set was derived from anesthesia billing records. Cardiac and nonprocedural anesthetic records were removed from our data set to conform more closely to the sample population used for derivation of the RQI with the goal of evaluating the generalizability of the RQI as a novel severity scoring methodology using primary CPT codes. The most common unmatched CPT codes listed in table 5 represent procedures for which an RQI score would be useful. A robust capture of CPT codes for 30-day mortality RQI calculation is important because the current analysis indicates these procedures were associated with a significantly higher ASA physical status and 30-day mortality.

There are a number of limitations that must be appreciated when using administrative data. Code definitions change with time. The RQI was derived using data from 2005 to 2008, whereas the current data timeframe was 2006–2011. Thus, only 3 yr overlap when comparing data sets. Review of the total number of CPT changes for the years 2005–2011 has been shown to total more than 2,500 changes.[26–32] Furthermore, the RQI incorporates the ASA physical status,[33,34] which has notable provider variability.[35,36] Regional or institutional differences in coding practices may also contribute to coding variability. Sources of coding error occur along the entire patient trajectory.[37–39] Error rates in coding have been shown to range from approximately 10–20%.[40,41]

Risk indices that use administrative data may not adequately assess how well hospitals and providers identify and respond to adverse events that developed during their care; described as "failure to rescue."‡‡ The degree to which in-hospital complications are documented differs among institutions. Rewarding hospitals solely for lower-documented in-hospital complications rates, such as using the RSIs for risk-adjusted comparison of outcomes among institutions, may not reflect important quality standards such as "monitoring" and "action taken."

Additionally, the current study has several limitations. The authors were unable to identify present-on-admission diagnoses and index procedures in order to exclude conditions or procedures that occurred during the hospitalization. As a result, including non–present-on-admission data may lead to overly optimistic predictive capability of the RSI in the current validation. As these data becomes more widely available, models may be used to assess risk of hospital-acquired conditions based on admission diagnoses or planned procedures.[19]

The authors' institution does not routinely collect the endpoints required for validating 30-day morbidity outside of participation in National Surgical Quality Improvement Project.

‡‡ Agency for Healthcare Research and Quality: Patient Safety Network Glossary, Failure to Rescue. Available at: http://psnet.ahrq.gov/popup_glossary.aspx?name=failuretorescue. Accessed February 2, 2013.

Table 6. AUROCs–RQI *vs.* Age + ASA, Age + PSS, ASA + PSS

	AUROC
RQI	0.888 (0.879–0.897)
ASA + PSS	0.880 (0.870–0.890)
Age + ASA	0.857 (0.847–0.868)
Age + PSS	0.842 (0.832–0.852)

ASA = American Society of Anesthesiologists physical status; AUROC = area under the receiver operating characteristic; PSS = Procedure Severity Score; RQI = Risk Quantification Index.

As a result, the authors were unable to assess the performance of the RQI for 30-day morbidity. In addition, the model for RQI 30-day mortality was derived using inpatient and outpatient surgical data. The current data set included inpatient surgeries only; thus, the authors were unable to validate RQI 30-day mortality model performance for outpatient surgery.

Conclusions

Although the RSI models for risk-adjusted healthcare outcomes demonstrated excellent discrimination, the poor calibration of the current models raises concerns about their generalizability. Assessment of calibration of the MEDPAR data set used to generate the original RSIs would be informative, with the potential to rederive the risk associated with covariates of interest to improve its performance on external data sets.

The RQI for 30-day mortality performed well on the current data set for matched CPT data. However, the current data reveal a large number of unmatched CPT codes for cases associated with significantly higher morbidity and mortality. A robust capture of CPT codes for 30-day mortality RQI calculation may identify patients at increased risk. Inclusion of age in the RQI was of limited additional predictive information in the analysis of the current data set.

The authors thank Frank E. Harrell, Jr., Ph.D., Chair and Professor of Biostatistics, Vanderbilt University, Nashville, Tennessee, for his invaluable assistance with statistical analysis.

References

1. Charlson ME, Pompei P, Ales KL, MacKenzie CR: A new method of classifying prognostic comorbidity in longitudinal studies: Development and validation. J Chronic Dis 1987; 40:373–83
2. Deyo RA, Cherkin DC, Ciol MA: Adapting a clinical comorbidity index for use with ICD-9-CM administrative databases. J Clin Epidemiol 1992; 45:613–9
3. Romano PS, Roos LL, Jollis JG: Adapting a clinical comorbidity index for use with ICD-9-CM administrative data: Differing perspectives. J Clin Epidemiol 1993; 46:1075–9; discussion 1081–90
4. van Walraven C, Austin PC, Jennings A, Quan H, Forster AJ: A modification of the Elixhauser comorbidity measures into a point system for hospital death using administrative data. Med Care 2009; 47:626–33
5. van Walraven C, Wong J, Bennett C, Forster AJ: The Procedural Index for Mortality Risk (PIMR): An index calculated using administrative data to quantify the independent influence of procedures on risk of hospital death. BMC Health Serv Res 2011; 11:258
6. Sessler DI, Sigl JC, Manberg PJ, Kelley SD, Schubert A, Chamoun NG: Broadly applicable risk stratification system for predicting duration of hospitalization and mortality. Anesthesiology 2010; 113:1026–37
7. Cohen NA, Hannenberg AA: Risk stratification index: An important advance in comparing health care apples to oranges. Anesthesiology 2010; 113:1004–6
8. Pace NL: Incomplete validation of risk stratification indices. Anesthesiology 2011; 115:214–5; author reply 215–6
9. Harrell FE: Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis. New York, Springer-Verlag, 2001, pp 493–4
10. Sigl JC, Sessler DI, Kelley SD, Chamoun NG: Incomplete Validation of Risk Stratification Indices. Anesthesiology 2011; 115:215–6
11. Dalton JE, Kurz A, Turan A, Mascha EJ, Sessler DI, Saager L: Development and validation of a risk quantification index for 30-day postoperative mortality and morbidity in noncardiac surgical patients. Anesthesiology 2011; 114:1336–44
12. American College of Surgeons National Surgical Quality Improvement Program: User Guide for the 2010 Participant Use Data File. Chicago, American College of Surgeons, 2010
13. Murphy SN, Chueh HC: A security architecture for query tools used to access large biomedical databases. Proc AMIA Symp 2002; 552–6
14. Murphy SN, Gainer V, Chueh HC: A visual interface designed for novice users to find research patient cohorts in a large biomedical database. AMIA Annu Symp Proc 2003;489–93
15. Centers for Disease Control and Prevention National Center for Health Statistics: National Death Index, Matching Criteria. Hyattsville, U.S. Department of Health and Human Services, 2011
16. Steyerberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, Obuchowski N, Pencina MJ, Kattan MW: Assessing the performance of prediction models: A framework for traditional and novel measures. Epidemiology 2010; 21:128–38
17. Department of Health and Human Services: 45 CFR Part 162, Proposed Rules. Federal Register, Vol. 77, No. 74, 2012
18. Agency for Healthcare Research and Quality Healthcare Cost and Utilization Project (HCUP): Clinical Classifications Software 2012. Healthcare Cost and Utilization Project—HCUP a Federal-State-Industry Partnership in Health Data, 2011
19. Cook NR: Use and misuse of the receiver operating characteristic curve in risk prediction. Circulation 2007; 115:928–35
20. Brier GW: Verification of forecasts expressed in terms of probability. Monthly Weather Review 1950; 78:1–3
21. Reistetter TA, Graham JE, Deutsch A, Markello SJ, Granger CV, Ottenbacher KJ: Diabetes comorbidity and age influence rehabilitation outcomes after hip fracture. Diabetes Care 2011; 34:1375–7
22. 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
23. Goldman L, Caldera DL, Nussbaum SR, Southwick FS, Krogstad D, Murray B, Burke DS, O'Malley TA, Goroll AH, Caplan CH, Nolan J, Carabello B, Slater EE: Multifactorial index of cardiac risk in noncardiac surgical procedures. N Engl J Med 1977; 297:845–50
24. Goldman L: The revised cardiac risk index delivers what it promised. Ann Intern Med 2010; 152:57–8
25. Raval MV, Cohen ME, Ingraham AM, Dimick JB, Osborne NH, Hamilton BH, Ko CY, Hall BL: Improving American College of Surgeons National Surgical Quality Improvement Program

risk adjustment: Incorporation of a novel procedure risk score. J Am Coll Surg 2010; 211:715–23
26. American Medical Association: Current Procedural Terminology, Standard Edition. Chicago, American Medical Association Press, 2005
27. American Medical Association: Current Procedural Terminology, Standard Edition. Chicago, American Medical Association Press, 2006
28. American Medical Association: Current Procedural Terminology, Standard Edition. Chicago, American Medical Association Press, 2007
29. American Medical Association: Current Procedural Terminology, Standard Edition. Chicago, American Medical Association Press, 2008
30. American Medical Association: Current Procedural Terminology, Standard Edition. Chicago, American Medical Association Press, 2009
31. American Medical Association: Current Procedural Terminology, Standard Edition. Chicago, American Medical Association Press, 2010
32. American Medical Association: Current Procedural Terminology, Standard Edition. Chicago, American Medical Association Press, 2011
33. Saklad M: Grading of patients for surgical procedures. Anesthesiology 1941; 2:281–4
34. American Society of Anesthesiologists: New classification of physical status. Anesthesiology 1963; 24:111
35. Haynes SR, Lawler PG: An assessment of the consistency of ASA physical status classification allocation. Anaesthesia 1995; 50:195–9
36. Aronson WL, McAuliffe MS, Miller K: Variability in the American Society of Anesthesiologists Physical Status Classification Scale. AANA J 2003; 71:265–74
37. O'Malley KJ, Cook KF, Price MD, Wildes KR, Hurdle JF, Ashton CM: Measuring diagnoses: ICD code accuracy. Health Serv Res 2005; 40(5 Pt 2):1620–39
38. Green J, Wintfeld N: How accurate are hospital discharge data for evaluating effectiveness of care? Med Care 1993; 31:719–31
39. Lloyd SS, Rissing JP: Physician and coding errors in patient records. JAMA 1985; 254:1330–6
40. Hsia DC, Krushat WM, Fagan AB, Tebbutt JA, Kusserow RP: Accuracy of diagnostic coding for Medicare patients under the prospective-payment system. N Engl J Med 1988; 318:352–5
41. Goldstein LB: Accuracy of ICD-9-CM coding for the identification of patients with acute ischemic stroke: Effect of modifier codes. Stroke 1998; 29:1602–4

Appendix 1. RSI In-hospital Mortality Predictors and Coefficients with Notation of "Self-fulfilling" Codes

Self-Fulfilling Code	Predictors	Description	ICD-9-CM Codes Included in Covariate	Coefficient (β)
0	D038	Septicemia	All subcodes except 0380, 03810, 03811, and 03819	0.438
0	D0380	Streptococcal septicemia	All subcodes	0.544
0	D03810	Staphylococcal septicemia, unspecified	None	0.883
0	D03811	Methicillin-susceptible *Staphylococcus aureus* septicemia	None	1.066
0	D03819	Other staphylococcal septicemia	None	0.629
0	D0388	Other specified septicemias	All subcodes	1.091
0	D0389	Unspecified septicemia	All subcodes	1.16
0	D157	Malignant neoplasm of pancreas	All subcodes	0.871
0	D1622	Malignant neoplasm of main bronchus	All subcodes	0.872
0	D1623	Malignant neoplasm of upper lobe, bronchus, or lung	All subcodes	0.405
0	D1628	Malignant neoplasm of other parts of bronchus or lung	All subcodes	1.036
0	D1629	Malignant neoplasm of bronchus and lung, unspecified	All subcodes	0.965
0	D1970	Secondary malignant neoplasm of lung	All subcodes	0.626
0	D1972	Secondary malignant neoplasm of pleura	All subcodes	0.865
0	D1976	Secondary malignant neoplasm of retroperitoneum and peritoneum	All subcodes	0.764
0	D1977	Malignant neoplasm of liver, secondary	All subcodes	0.911
0	D1978	Secondary malignant neoplasm of other digestive organs and spleen	All subcodes	0.608
0	D198	Secondary malignant neoplasm of other specified sites	All subcodes except 1983 and 1985	0.44
0	D1983	Secondary malignant neoplasm of brain and spinal cord	All subcodes	0.604

(*Continued*)

Appendix 1. (*Continued*)

Self-Fulfilling Code	Predictors	Description	ICD-9-CM Codes Included in Covariate	Coefficient (β)
0	D1985	Secondary malignant neoplasm of bone and bone marrow	All subcodes	0.774
0	D199	Malignant neoplasm without specification of site	All subcodes except 1991	1.486
0	D1991	Other malignant neoplasm without specification of site	All subcodes	0.896
0	D200	Lymphosarcoma and reticulosarcoma and other specified malignant tumors of lymphatic tissue	All subcodes	0.905
0	D2028	Other lymphomas	All subcodes except 20280 and 20283	0.878
0	D20280	Other malignant lymphomas, unspecified site, extranodal and solid organ sites	None	0.629
0	D20283	Other malignant lymphomas, intraabdominal lymph nodes	None	0.849
0	D20300	Multiple myeloma, without mention of having achieved remission	None	0.655
0	D20410	Chronic lymphoid leukemia, without mention of having achieved remission	None	0.525
0	D20500	Acute myeloid leukemia, without mention of having achieved remission	None	1.759
0	D208	Leukemia of unspecified cell type	All subcodes	1.26
0	D2113	Benign neoplasm of colon	All subcodes	−0.882
0	D2387	Other lymphatic and hematopoietic tissues	All subcodes	0.473
0	D261	Nutritional marasmus	All subcodes	0.96
0	D262	Other severe protein-calorie malnutrition	All subcodes	0.978
0	D2639	Unspecified protein-calorie malnutrition	All subcodes	0.56
0	D2720	Pure hypercholesterolemia	All subcodes	−0.578
0	D2724	Other and unspecified hyperlipidemia	All subcodes	−0.569
0	D27542	Hypercalcemia	None	0.554
0	D2760	Hyperosmolality and/or hypernatremia	All subcodes	0.5
0	D2761	Hyposmolality and/or hyponatremia	All subcodes	0.207
0	D2762	Acidosis	All subcodes	0.648
0	D2765	Volume depletion	All subcodes	0.464
0	D2767	Hyperpotassemia	All subcodes	0.346
0	D2768	Hypopotassemia	All subcodes	−0.172
0	D2780	Overweight and obesity	All subcodes	−0.498
0	D2809	Iron deficiency anemia, unspecified	All subcodes	−0.377
0	D2866	Defibrination syndrome	All subcodes	1.527
0	D2869	Other and unspecified coagulation defects	All subcodes	0.559
0	D2875	Thrombocytopenia, unspecified	All subcodes	0.331
0	D2900	Senile dementia, uncomplicated	All subcodes	0.513
0	D2948	Other persistent mental disorders due to conditions classified elsewhere	All subcodes	0.369
0	D3051	Tobacco use disorder	All subcodes	−0.367
0	D348	Cerebral cysts	All subcodes except 3481 and 3483	1.142
1	D3481	Anoxic brain damage	All subcodes	1.919
1	D3483	Encephalopathy, not elsewhere classified	All subcodes	0.653
0	D4019	Unspecified essential hypertension	All subcodes	−0.261

(*Continued*)

Appendix 1. (*Continued*)

Self-Fulfilling Code	Predictors	Description	ICD-9-CM Codes Included in Covariate	Coefficient (β)
0	D40391	Hypertensive chronic kidney disease, unspecified, with chronic kidney disease stage V or end-stage renal disease	None	0.135
1	D41001	Acute myocardial infarction of anterolateral wall, initial episode of care	None	1.47
1	D41011	Acute myocardial infarction of other anterior wall, initial episode of care	None	1.355
1	D41021	Acute myocardial infarction of inferolateral wall, initial episode of care	None	1.377
1	D41041	Acute myocardial infarction of other inferior wall, initial episode of care	None	1.172
1	D41071	Subendocardial infarction, initial episode of care	None	0.564
1	D41081	Acute myocardial infarction of other specified sites, initial episode of care	None	1.227
1	D41091	Acute myocardial infarction of unspecified site, initial episode of care	None	1.56
0	D4139	Other and unspecified angina pectoris	All subcodes	−0.209
1	D41511	Iatrogenic pulmonary embolism and infarction	None	1.426
1	D41519	Other pulmonary embolism and infarction	None	0.868
0	D4241	Aortic valve disorders	All subcodes	0.205
1	D427	Cardiac dysrhythmias	All subcodes except 42731 and 42741	0.647
0	D42731	Atrial fibrillation	None	0.321
1	D42741	Ventricular fibrillation	None	1.974
1	D428	Heart failure	All subcodes	0.519
1	D430	Subarachnoid hemorrhage	All subcodes	2.334
1	D431	Intracerebral hemorrhage	All subcodes	2.639
1	D432	Other and unspecified intracranial hemorrhage	All subcodes except 4321	2.226
1	D4321	Subdural hemorrhage	All subcodes	1.211
1	D4340	Cerebral thrombosis	All subcodes	1.419
1	D43411	Cerebral embolism with cerebral infarction	None	1.059
1	D43491	Cerebral artery occlusion, unspecified with cerebral infarction	None	1.189
0	D4359	Unspecified transient cerebral ischemia	All subcodes	−0.555
1	D436	Acute, but ill-defined, cerebrovascular disease	All subcodes	1.072
1	D43820	Late effects of cerebrovascular disease, hemiplegia affecting unspecified side	None	0.262
0	D44021	Atherosclerosis of native arteries of the extremities with intermittent claudication	None	−0.589
1	D44024	Atherosclerosis of native arteries of the extremities with gangrene	None	0.656
1	D441	Aortic aneurysm and dissection	All subcodes except 4410 and 4413	0.283
1	D4410	Dissection of aorta	All subcodes	1.996
1	D4413	Abdominal aneurysm, ruptured	All subcodes	2.982
1	D44422	Arterial embolism and thrombosis of lower extremity	None	0.844
0	D4550	Internal hemorrhoids without mention of complication	All subcodes	−0.926
1	D4582	Iatrogenic hypotension	All subcodes	0.476

(*Continued*)

Appendix 1. (*Continued*)

Self-Fulfilling Code	Predictors	Description	ICD-9-CM Codes Included in Covariate	Coefficient (β)
1	D4589	Hypotension, unspecified	All subcodes	0.795
0	D4824	Pneumonia due to *Staphylococcus*	All subcodes	0.471
0	D485	Bronchopneumonia, organism unspecified	All subcodes	1.122
0	D486	Pneumonia, organism unspecified	All subcodes	0.705
0	D493	Asthma	All subcodes	−0.331
0	D496	Chronic airway obstruction, not elsewhere classified	All subcodes	0.18
0	D5070	Pneumonitis due to inhalation of food or vomitus	All subcodes	1.149
0	D5119	Unspecified pleural effusion	All subcodes	0.411
1	D5128	Other spontaneous pneumothorax	All subcodes	0.62
0	D515	Postinflammatory pulmonary fibrosis	All subcodes	0.519
0	D5163	Idiopathic fibrosing alveolitis	All subcodes	0.97
1	D5180	Pulmonary collapse	All subcodes	−0.212
1	D5185	Pulmonary insufficiency following trauma and surgery	All subcodes	1.121
0	D5188	Other diseases of lung	All subcodes	1.095
0	D5533	Diaphragmatic hernia without mention of obstruction or gangrene	All subcodes	−0.358
1	D5570	Acute vascular insufficiency of intestine	All subcodes	1.554
1	D5579	Unspecified vascular insufficiency of intestine	All subcodes	0.72
1	D5609	Unspecified intestinal obstruction	All subcodes	0.487
1	D567	Peritonitis and retroperitoneal infections	All subcodes	0.816
1	D570	Acute and subacute necrosis of liver	All subcodes	1.342
0	D5722	Hepatic encephalopathy	All subcodes	1.254
0	D5728	Other sequelae of chronic liver disease	All subcodes	1.596
1	D5780	Hematemesis	All subcodes	0.764
1	D5789	Hemorrhage of gastrointestinal tract, unspecified	All subcodes	0.575
0	D5845	Acute kidney failure with lesion of tubular necrosis	All subcodes	0.809
0	D5849	Acute renal failure, unspecified	All subcodes	0.771
0	D585	Chronic kidney disease	All subcodes	0.302
0	D586	Renal failure, unspecified	All subcodes	1.099
0	D5939	Unspecified disorder of kidney and ureter	All subcodes	0.205
0	D7070	Chronic ulcer of skin	All subcodes	0.609
0	D71595	Osteoarthrosis, unspecified whether generalized or localized, pelvic region and thigh	None	−1.044
0	D7169	Arthropathy, unspecified	All subcodes	−0.358
1	D78001	Coma	None	2.174
0	D7812	Abnormality of gait	All subcodes	−0.87
0	D783	Symptoms concerning nutrition, metabolism, and development	All subcodes	0.429
0	D7840	Headache	All subcodes	−0.777
0	D7854	Gangrene	All subcodes	0.661
1	D7855	Shock without mention of trauma	All subcodes	1.299
0	D7860	Enlargement of lymph nodes	All subcodes	0.252
0	D78820	Retention of urine, unspecified	None	−0.27
0	D789	Other symptoms involving abdomen and pelvis	All subcodes	0.293
1	D799	Other ill-defined and unknown causes of morbidity and mortality	All subcodes	0.672

(*Continued*)

Appendix 1. (*Continued*)

Self-Fulfilling Code	Predictors	Description	ICD-9-CM Codes Included in Covariate	Coefficient (β)
1	D801	Fracture of base of skull	All subcodes	1.618
0	D8050	Cervical closed fracture of vertebral column without mention of spinal cord injury	All subcodes	1.064
1	D851	Cerebral laceration and contusion	All subcodes	1.359
1	D852	Subarachnoid, subdural, and extradural hemorrhage, after injury	All subcodes	1.577
1	D853	Other and unspecified intracranial hemorrhage after injury	All subcodes	1.711
1	D99702	Iatrogenic cerebrovascular infarction or hemorrhage	None	1.506
1	D99811	Hemorrhage complicating a procedure	None	0.569
0	DV46	Encounter for dialysis and dialysis catheter care	All subcodes	0.395
0	DV5789	Care involving other specified rehabilitation procedure	None	−0.846
0	DV58	Encounter for other and unspecified procedures and aftercare	All subcodes	−0.323
1	DV66	Convalescence and palliative care	All subcodes	2.614
1	P0051	Implantation of cardiac resynchronization defibrillator, total system	None	−1.666
1	P0066	Percutaneous transluminal coronary angioplasty or coronary atherectomy	None	−0.792
0	P0309	Other exploration and decompression of spinal canal	None	−1.085
0	P0331	Spinal tap	None	0.365
0	P3404	Insertion of intercostal catheter for drainage	None	0.515
1	P3601	Single vessel percutaneous transluminal coronary angioplasty or coronary atherectomy without mention of thrombolytic agent	None	−0.528
1	P370	Pericardiocentesis	All subcodes	1.141
0	P3726	Catheter based invasive electrophysiologic testing	None	−1.224
1	P3761	Implant of pulsation balloon	None	1.674
1	P3778	Insertion of temporary transvenous pacemaker system	None	0.572
1	P3794	Implantation or replacement of automatic cardioverter/ defibrillator, total system	None	−1.826
0	P3812	Endarterectomy, other vessels of head and neck	None	−0.981
0	P3893	Venous catheterization, not elsewhere classified	None	0.651
0	P3895	Venous catheterization for renal dialysis	None	0.401
0	P3898	Other puncture of artery	None	0.618
0	P4516	Esophagogastroduodenoscopy with closed biopsy	None	−0.362
0	P4523	Colonoscopy	None	−0.538
0	P5123	Laparoscopic cholecystectomy	None	−0.779
0	P5411	Exploratory laparotomy	None	1.19
0	P5491	Percutaneous abdominal drainage	None	0.479
0	P5794	Insertion of indwelling urinary catheter	None	0.454
0	P6029	Other transurethral prostatectomy	None	−1.204
0	P8154	Total knee replacement	None	−2.033
0	P8192	Injection of therapeutic substance into oint or ligament	None	−1.033
0	P8411	Amputation of toe	None	−0.622

(*Continued*)

Appendix 1. (*Continued*)

Self-Fulfilling Code	Predictors	Description	ICD-9-CM Codes Included in Covariate	Coefficient (β)
0	P8853	Angiocardiography of left heart structures	None	−0.624
0	P8872	Diagnostic ultrasound of heart	None	−0.305
0	P8891	Magnetic resonance imaging of brain and brain stem	None	−0.486
0	P8964	Pulmonary artery wedge monitoring	None	0.606
0	P8965	Measurement of systemic arterial blood gases	None	0.646
0	P9205	Cardiovascular and hematopoietic scan and radioisotope function study	None	−1.024
0	P9390	Noninvasive mechanical ventilation	None	0.932
0	P9394	Respiratory medication administered by nebulizer	None	0.529
0	P9604	Insertion of endotracheal tube	None	1.244
0	P9605	Other intubation of respiratory tract	None	0.928
0	P9607	Insertion of other (naso-)gastric tube	None	0.573
1	P9671	Continuous invasive mechanical ventilation for <96 consecutive hours	None	0.698
0	P9905	Transfusion of platelets	None	0.701
0	P9907	Transfusion of other serum	None	0.526
0	P9915	Parenteral infusion of concentrated nutritional substances	None	0.648
1	P9960	Cardiopulmonary resuscitation, not otherwise specified	None	4.057
	Constant	Constant		−2.198

ICD-9-CM = International Classification of Disease, Ninth Revision, Clinical Modification; RSI = Risk Stratification Index.

Appendix 2. Revised Code for Calculating RQI 30-day Mortality

```
RQI.mortality.logOdds = –16.272363
  +0.10429722 * pss.mortality
  +0.026466389 * age
  –5.6279714e-06 * max(age - 25, 0)^3
  +3.8473642e-05 * max(age - 48, 0)^3
  –5.1963656e-05 * max(age - 63, 0)^3
  +1.9117985e-05 * max(age - 82, 0)^3
  +0.76745424 * (asaclass = "II")
  +2.1582119 * (asaclass = "III")
  +3.5228968 * (asaclass = "IV")
  +4.9336867 * (asaclass = "V")
```

Revised code for calculating RQI 30-day mortality. Jarrod Dalton, M.A., Senior Biostatistician, Departments of Quantitative Health Sciences and Outcomes Research, Cleveland Clinic, Cleveland, Ohio. Electronic communication, June 11th, 2012.

asaclass = American Society of Anesthesiologists Physical Status; max = maximum; pss = Procedure Severity Score; RQI = Risk Quantification Index.

Appendix 3. SPSS Code Used to Calculate RQI 30-day Mortality

```
COMPUTE ASAFactor=0.
If (ASAClass = 2) ASAFactor=7.
If (ASAClass = 3) ASAFactor=20.
If (ASAClass = 4) ASAFactor=35.
If (ASAClass = 5) ASAFactor=48.
COMPUTE AgeFactor=(Age-10)/90*24.
COMPUTE SumFactor=ASAFactor+AgeFactor+(PSSMortality).
EXECUTE.
```

SPSS code used to calculate RQI 30-d mortality. Jonathan P. Wanderer, M.D., M.Phil., Instructor, Department of Anesthesia, Vanderbilt University, Nashville, Tennessee. Last modified December 25, 2012.

ASA = American Society of Anesthesiologists Physical Status; PSS = Procedure Severity Score; RQI = Risk Quantification Index.

Appendix 4. R Code for RSI and RQI Calibration Curves

```
# Generate calibration curves for RSI/RQI data
##
# This code was run using R version 2.15.1 and package rms version 3.5-0
#
#
# Jon Wanderer. Last modified 12/25/2012.
#
# First we load the necessary libraries
library(rms)
library(foreign)
# Then we load our data from our SPSS data file
rsi = read.spss('C:\\Run Hard\\Calibration\
\RSI_reverse_Logit_calibration.ALLVARIABLES_2.sav')
# Next we generate the calibration curves for RSI
val.prob(, rsi$DeadAtDischarge, rsi$RSI_INHOSP_ADJUST, m=500, legendloc=FALSE)
val.prob(, rsi$DeadAt1MO, rsi$RSI_30Day, m=500, legendloc=FALSE)
val.prob(, rsi$DeadAt1YR, rsi$RSI_1YR, m=500, legendloc=FALSE)
val.prob(, rsi$MedianLOSOrBelow, rsi$RSI_LOS, m=500)
# Then we load our data for the RQI
rqi = read.spss('C:\\Run Hard\\Calibration\\rqi redo deidentified.sav')
# Finally we generate the calibration curve for the RQI
rqi_logit = rqi$RQIMortality
rqi_prob = rqi$PredMortality
rqi_outcome = rqi$DeadAt1MO
val.prob(rqi_prob, rqi_outcome, pl=TRUE, smooth=FALSE)
```

R code for RSI and RQI calibration curves. Jonathan P. Wanderer, M.D., M.Phil., Instructor, Department of Anesthesia, Vanderbilt University, Nashville, Tennessee. Last modified December 25, 2012.

RQI = Risk Quantification Index; RSI = Risk Stratification Index.