Comparison of an Updated Risk Stratification Index to Hierarchical Condition Categories

George F. Chamoun, B.S., Linyan Li, Sc.D., Nassib G. Chamoun, M.S., Vikas Saini, M.D., Daniel I. Sessler, M.D.

ABSTRACT

Background: The Risk Stratification Index and the Hierarchical Condition Categories model baseline risk using comorbidities and procedures. The Hierarchical Condition categories are rederived yearly, whereas the Risk Stratification Index has not been rederived since 2010. The two models have yet to be directly compared. The authors thus rederived the Risk Stratification Index using recent data and compared their results to contemporaneous Hierarchical Condition Categories.

Methods: The authors reimplemented procedures used to derive the original Risk Stratification Index derivation using the 2007 to 2011 Medicare Analysis and Provider review file. The Hierarchical Condition Categories were constructed on the entire data set using software provided by the Center for Medicare and Medicaid Services. C-Statistics were used to compare discrimination between the models. After calibration, accuracy for each model was evaluated by plotting observed against predicted event rates.

Results: Discrimination of the Risk Stratification Index improved after rederivation. The Risk Stratification Index discriminated considerably better than the Hierarchical Condition Categories for in-hospital, 30-day, and 1-yr mortality and for hospital length-of-stay. Calibration plots for both models demonstrated linear predictive accuracy, but the Risk Stratification Index predictions had less variance.

Conclusions: Risk Stratification discrimination and minimum-variance predictions make it superior to Hierarchical Condition Categories. The Risk Stratification Index provides a solid basis for care-quality metrics and for provider comparisons.
(ANESTHESIOLOGY 2018; 128:109-16)

H OSPITAL performance metrics are increasingly available to patients and are linked to financial consequences for physicians and institutions. Because baseline and procedural risk varies by institution, accurate risk adjustment models are necessary to fairly compare outcomes across health delivery systems. Commonly used risk adjustment systems include the Charlson Comorbidity Index,[1] the Elixhauser Comorbidity Index,[2] and the American Society of Anesthesiologists physical status score.[3] Others include the Procedural Index for Mortality Risk,[4] the Risk Quantification Index,[5] the Preoperative Score to Predict Postoperative Mortality,[6] and the Surgical Mortality and Probability Model.[7] Few are fully validated on independent populations.

The Risk Stratification Index (RSI) is a broadly applicable risk adjustment measure for predicting mortality across various time horizons and hospital length-of-stay.[8] RSI models have been developed for inpatient, 30-day, and 1-yr mortality, as well as for length-of-stay. The models were derived from more than 35 million Medicare hospitalizations between 2001 to 2006 and were thereafter validated across the entire range of adult ages in California inpatients[9] and in three single-center studies.[8,10,11] RSI models additionally demonstrated superior discrimination when compared to the Charlson Comorbidity Index.[8] More recently, the RSI models performed well in a new set of 39 million Medicare admissions from 2008 to 2012.[12]

What We Already Know about This Topic

- Accurate risk adjustment models are needed to compare outcomes across healthcare delivery systems
- The Risk Stratification Index and the Hierarchical Condition Categories model baseline risk using comorbidities and procedures
- The Hierarchical Condition categories are rederived yearly, whereas the Risk Stratification Index has not been rederived since 2010

What This Study Tells Us That Is New

- Discrimination of the Risk Stratification Index improved after rederivation
- The Risk Stratification Index discriminated considerably better than the Hierarchical Condition Categories
- Calibration plots for both models demonstrated linear predictive accuracy, but the Risk Stratification Index predictions had less variance

Nonetheless, published RSI models[8] date to 2010, and there have since been important changes. For example, the 2001 to 2006 Medicare provider analysis and review (MEDPAR) data set was restricted to nine diagnostic and six procedure codes. But since 2010, Medicare has allowed up to 25 diagnostic codes and 25 procedure codes per admission. Furthermore, until 2008, Medicare participants were not required to identify conditions as being present-on-admission.[13] The problem with including

diagnostic codes consequent to hospital-acquired complications is that the complications are attributed to baseline illness, thus improving apparent performance. Our initial goal was thus to update the original RSI coefficients (version 1.0) using recent Medicare data, a suite of models we refer to as RSI 2.0.

Another risk adjustment model is the Hierarchical Condition Category (HCC) community score which is widely used to adjust for comorbidities related to cost and overall risk.[14] This single overall risk metric encompasses mortality at various postadmission times, combined with expected cost. While having the benefit of simplicity, the single metric may poorly characterize various outcomes. We therefore compared performance of the Risk Stratification Index to the HCC.

Materials and Methods

Under a Data Use Agreement with the Center for Medicare and Medicaid services, more than 98,077,959 patient-stay records from the 2007 to 2014 MEDPAR files served as our complete data set. The MEDPAR file contains claims data from Medicare-certified inpatient hospitals and skilled nursing facilities. We selected the Medicare population due to its national representation, size, high baseline risk, diverse set of beneficiaries, and broad set of potential risk strata. Each record in the file corresponds to a qualified hospital stay. Fields include up to 25 diagnostic and 25 procedure International Classification of Diseases 9th Revision (ICD-9) codes as well as age, sex, race, provider information, and an encrypted beneficiary identification key.

Stays without at least one procedure performed or where the beneficiary was less than 65 yr of age were excluded from analysis. When building our models, we withheld records from the 2012 to 2014 files from the final analysis data set to add true independence to our validations. We divided our 2007 to 2011 file into development and validation cohorts using unique beneficiary identification codes (split sample, 50% development, 50% validation). We randomly selected

development and validation care episodes by beneficiary rather than by hospital to define genuine cohorts.

When validating our models on present-on-admission (POA) data, we restricted analysis to hospitalizations for which POA coding was available. For the POA model, we only considered procedures that fell into the same Clinical Classification Software[15] category as the primary procedure. We were thus able to incorporate procedural risk without giving hospitals credit for procedures linked to care complications. Finally, we computed a comorbidities only model by withholding all ICD-9 procedure codes from analysis. This was to additionally ensure a fair comparison against HCC, a comorbidities only model, and ensure that our models were not surgically slanted. Table 1 summarizes each of the data sets and their respective exclusions.

Model Development

Statistical procedures were performed using SPSS Statistics version 22.0 (IBM, USA) and SAS version 9.4 (SAS, USA). Our previous ICD-9 code stratification algorithm[8] was reimplemented with minor modifications. Specifically, when selecting diagnostic and procedure codes on the final iteration, we kept codes with an average annual incidence of at least 1,000 instead of the higher thresholds used in the original revision. We used the final analysis data set (table 1) to build our risk strata.

Cox regression was used to model postdischarge mortality and length of stay, whereas in-hospital mortality was modeled using logistic regression. For each endpoint, a Cox or logistic model was performed, and variables below a certain significance threshold were eliminated. For computational efficiency, random samples of increasing size were progressively taken from the development cohort, and variables were selected in a stepwise manner from the regressions using decreasing levels of significance with respect to the sample size. Variables eliminated in earlier regressions were not included in later iterations. This process is detailed in our original RSI publication.[8] The

Table 1. Summary of Data Sets

Data Set	N	Exclusion 1	Exclusion 2	Exclusion 3	Exclusion 4	Used For
All years MEDPAR file	98,077,959					Base file
Final analysis data set	33,101,922	Nonprocedure stays	Age < 65 yr	Years 2012–2014		ICD-9 code stratification*
Development cohort	16,551,203	Nonprocedure stays	Age < 65 yr	Years 2012–2014	50% random sample	Regression models
Validation cohort	16,550,719	Nonprocedure stays	Age < 65 yr	Years 2012–2014	50% random sample	Model validation
Prospective 2012–2014 data set	19,817,421	Nonprocedure stays	Age < 65 yr	Years 2007–2011		Independent validation† and comorbidities-only validation
Prospective POA data set	19,357,629	Nonprocedure stays	Age < 65 yr	Years 2007–2011	Non-POA compliant cases‡	Present-on-admission validation

* Stratification is the process by which regression risk classes are derived. In this case, we counted all ICD-9 codes in the given data set. †Withheld from all derivational procedures. ‡Years 2007 to mid 2008 were not compliant. There are also certain cases sprinkled throughout all years that are noncompliant. MEDPAR = Medicare provider analysis and review; POA = present-on-admission.

RSI 2.0 model coefficients were then applied prospectively to the validation cohort, as well as to the years 2012 to 2014.

We computed the HCC Community score for each stay record in our data set, using the yearly software publicly available at the Centers for Medicare and Medicaid Services website (https://www.cms.gov/Medicare/Health-Plans/MedicareAdvtgSpecRateStats/Risk-Adjustors.html; accessed September 21, 2017). Because HCC already includes demographic characteristics, we did not add them to the reported models. However, we computed additional C-statistics for RSI including demographics, as well as C-statistics, for demographics alone.

Our models were calibrated to enhance usability. Our calibration procedures aligned observed outcomes with predicted values using cumulative survival distributions as previously presented.[12] Calibration function parameters were estimated on the development data set using a nonlinear least squares algorithm, with logistic and extreme value functions serving as the foundations for our calibration functions.

Statistical Comparisons

Discrimination was evaluated for all endpoints using the area under the receiver operating characteristics (ROC) curve, known as the C-statistic. For purposes of comparison, we similarly derived predictions using the original 2010 RSI model on the same data sets. ROCs were plotted for HCC, as well as against all RSI models. We computed nonparametric Spearman rank-correlation coefficients for each of the RSI 2.0 models and for the HCC model to evaluate collinearity between the two predictive metrics modeling the same outcome.

Rather than assess information criteria of our models, we simply chose to test RSI sampling variance in a bootstrapped analysis of decreasing sample size. We performed this analysis for HCC as well, to serve as a benchmark for comparison to demonstrate that our models are not overfit.

Calibration for all endpoints was demonstrated using plots of the observed event rates over the expected event rates across risk bins of 0.1% resolution. We similarly plotted HCC predictions against observed outcomes after applying the same calibration method used for RSI. Correlation between predicted and observed mortalities and the number of bins of width 0.001 having at least 500 observations served as our metrics of comparison between the two models.

Results

The prospective 2012 to 2014 data set, excluded completely from all derivational procedures, differed somewhat from previous MEDPAR years. For example, in the 2012 to 2014 data set, there are an average of four additional diagnostic codes per admission, length-of-stay is reduced by one day at the 75th percentile, and 1-yr mortality was 1.1% lower. Table 2 shows the full characteristics for each data set, as well as for the validation data set used in the original RSI 1.0 report.

We retained 1,654 diagnostic strata and 841 procedure strata for regression analysis after performing the truncation

algorithm on the 2007 to 2011 data set. Our risk strata hierarchically encapsulate 50% of the 22,000 total ICD-9 codes and 66% of the 17,000 codes that appeared at least once per yr between 2007 and 2011. Each of the RSI 2.0 endpoint-specific models ended up with a different combination of codes from this list as a result of the stepwise algorithm. Each model and its respective list of codes can be found on the Lown Institute website (https://lowninstitute.org/ResearchActivities/RSI/DataDownloads; accessed September 21, 2017).

Discrimination for the newly derived RSI 2.0 mortality and length-of-stay endpoints improved compared to the original RSI report by ~1% across all endpoints. When applying RSI models to the independent 2012 to 2014 validation data set, discrimination increased by an additional 1% for both RSI 1.0 and 2.0 models. All models performed better on more recent data.

When diagnostic codes were restricted to those POA, discrimination of RSI 2.0 declined more for certain endpoints than for others. The models that considered events postdischarge (1-yr and 30-day mortality) declined by an average of 2% points, whereas the models that considered events that took place during the qualified stay (length-of-stay and in-hospital mortality) declined by ~10%. These results are consistent with our previous report using the RSI 1.0 model coefficients.[9,12] When RSI models were restricted to using comorbidities only, we saw a small decline in accuracy with all models declining by ~1% with the exception of length-of-stay that declined close to 3.5%.

Adding demographic characteristics to RSI 2.0 only trivially improved overall model discrimination (ROC increased by a maximum of 0.02). Demographic characteristics alone delivered a maximum C-statistic of approximately 0.63 between all endpoints.

HCC models and RSI models were correlated. The nonparametric correlations for HCC and RSI were strongest for 1-yr mortality, with a Spearman's ρ value of 0.723, and weakest for length of stay with a Spearman's ρ value of –0.5.

HCC models had lower C-statistics than those of RSI 2.0 for all endpoints and for the RSI comorbidities-only model (tables 3 and 4). ROC curves demonstrate the differences between HCC performance and RSI performance in this respect (figs. 1–4).*

Bootstrapping of RSI demonstrated that despite many degrees of freedom, model stability is not compromised in samples of decreasing size. As can be noted from figure 5, RSI variance increases at a rate approximately equal to that of HCC with respect to sample size, on par with Gaussian sampling error, indicative of a model that is not overfit.

All models, both HCC and RSI, responded well to calibration (figs. 1–5). Observed *versus* expected plots for mortality and length of stay demonstrated a strong linear correlation

---

*The presented receiver operating characteristics curves for Risk Stratification Index and were built using the entire Medicare provider analysis and review (MEDPAR) file, from 2007 to 2014, excluding patients under 65 and those without a procedure.

Table 2. Characterization

Observations	Development Cohort 2007–2011 (N = 16,551,203)	Validation Cohort 2007–2011 (N = 16,550,719)	Prospective Present-on-admission 2012–2014 (N = 19,357,629)	Prospective Data Set 2012–2014 (N = 19,817,421)	Original RSI Validation Data Set 2001–2006 (N = 17,589,824)*	Final Analysis Data Set 2007–2011 (N = 33,101,922)
Age, yr	77±8	77±8	77±8	77±8	74.1±10.2	77±8
Female, %	55.4	55.4	54.7	54.1	54.4	55.3
White/black/other	84.4/10.3/5.3	84.3/10.3/5.4	83.4/10.6/6.8	83.4/10.7/5.9	82.5/12.2/5.4	84.3/10.4/5.3
Number of diagnostic codes	8.9±3.6	9.6±4.4	13.1±6.0	13.7±6.2	7.1±2.3	9.6±4.4
Number of procedure codes	2.6±1.9	2.6±1.9	2.8±2.4	2.8±2.5	2.6±1.7	2.6±1.9
Median length of stay [IQR]	4 [3, 8]	4 [3, 8]	4 [3, 7]	4 [3, 7]	5 [3, 8]	4 [3, 8]
In-hospital mortality, %	4.9	4.9	4.4	4.5	5.3	4.9
30-day postdischarge mortality, %	4.9	4.9	5.3	5.3	5	4.9
1-yr postdischarge mortality, %	23.2	23.1	24.1	24.2	19.3	23.1

*From original RSI publication.

IQR = interquartile range, the middle 75% of the data; RSI = Risk Stratification Index.

between event rates (figs. 6–10). When comparing the goodness-of-fit between the RSI models and HCC 1-yr mortality models, the RSI model had an R^2 of 0.999 for 974 data bins, and the HCC model had an R^2 of 0.998 for 781 bins using prediction bins of size 0.001 having at least 500 individuals.

Discussion

Overall the rederived RSI 2.0 models improved discrimination slightly. Thus, updating the models by recalculating the coefficients to reflect recent substantive changes in coding techniques only modestly enhanced performance. The original 2010 version and the current 2.0 version are therefore far more notable for their similarities than differences, suggesting that our derivation techniques are both robust and stable.

HCC is a multivariable model that includes just 76 comorbidities and no procedural information. It is derived using professional expertise and judgement; consequently, developers are able to identify coding categories that are highly discretionary and could cause coding variation bias or create coding incentives.[14] In contrast, RSI models are derived algorithmically and use only frequency and significance to determine which of approximately 30,000 potential ICD-9 codes will be included in each. In practice, between 450 and 1,400 codes and truncated risk-classes were included in our various models. RSI models thus include far more predictive information. It is thus unsurprising that RSI 2.0 discriminated better than HCC, thus validating our hypothesis and indicating that using RSI predicts better than HCC for various mortality horizons and duration of hospitalization.

One potential criticism of prediction models that include hundreds of variables is that they are potentially overfit. However, RSI models are not overfit because they were derived from enormous data sets, as evidenced by remarkably stable performance characteristics across years and data sets. Furthermore, all of the variables used in the RSI model have univariate P values less than 10^{-6}, and many have P values less than 10^{-300}. Our bootstrapping analysis further validated this claim. An additional criticism of models with large numbers of variables, such as RSI, is that they are harder to use than models with fewer variables such as HCC. But in practice, RSI and HCC are not used by clinicians for individual patients, and both are easily calculated with standard computer algorithms.

In addition to enhanced discrimination, using a large number of variables in model building enhances precision for probabilistic endpoint predictions. For example in figures 1 and 2, although calibrated RSI has a similarly high R^2 to HCC, HCC has 781 bins of prediction, whereas RSI has 974. This enhanced range results in a more significant correlation for RSI predictions than those of HCC. The additional variables included in RSI therefore appear to capture a substantial amount of risk overlooked by other models, diminishing variance between observed and expected event rates.

In addition to RSI and HCC, various models predicting perioperative outcomes have been proposed in recent years,

Table 3. Risk Stratification Index 1.0 Receiver Operating Characteristic All Data Sets

Endpoint	Development Data Set	Validation Data Set	POA Validation Data Set 2012–2014	Prospective Data Set 2012–2014
LOS	0.852	0.852	0.781	0.849
In-hospital	0.933	0.933	0.868	0.945
1-yr mortality	0.831	0.831	0.823	0.833
30-day mortality	0.835	0.836	0.821	0.843

95% CI for each value is ≤ 0.001.
LOS = length of stay; POA = present-on-admission.

Table 4. Risk Stratification Index 2.0 and HCC Receiver Operating Characteristics All Data Sets

Endpoint	Development Data Set	Validation Data Set	POA Validation Data Set 2012–2014	Prospective Data Set 2012–2014	Prospective Data Set 2012–2014 (Comorbidities-only)
Length of stay	0.857	0.857	0.787	0.857	0.823
In-hospital mortality	0.943	0.943	0.884	0.956	0.944
1-yr mortality	0.837	0.837	0.834	0.842	0.831
30-day mortality	0.847	0.847	0.840	0.862	0.851
HCC (1 yr)	0.783	0.783	0.793	0.793	0.793
HCC (in-hospital)	0.838	0.839	0.791	0.844	0.844

95% CI for all estimates ≤ 0.001.
HCC = Hierarchical Condition Category; POA = present-on-admission.
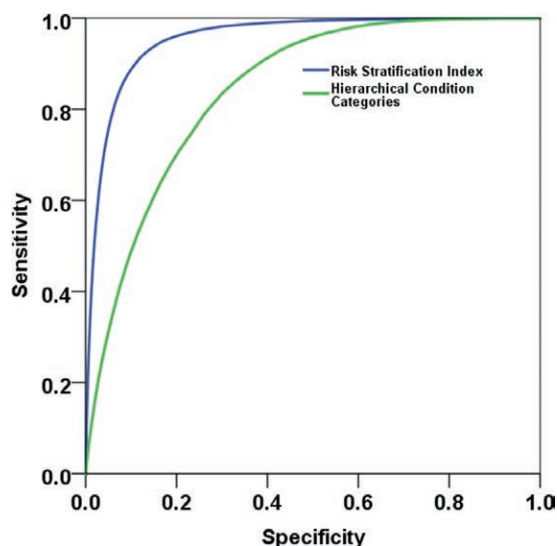
Fig. 1. Prediction of in-hospital mortality by Risk Stratification Index *versus* Hierarchical Condition Categories. The data are from a 5% sample of Medicare Analysis and Provider Review data set years 2007 to 2014.
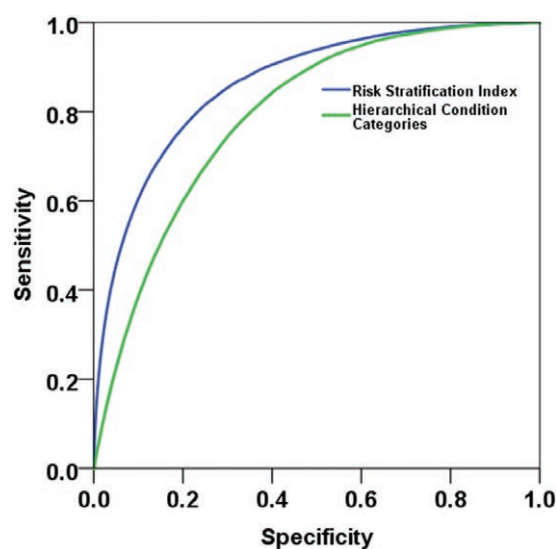
Fig. 2. Prediction of 30-day mortality by Risk Stratification Index *versus* Hierarchical Condition Categories. The data are from a 5% sample of Medicare Analysis and Provider Review data set years 2012 to 2014.

with various degrees of validation and calibration.[1–7] Many report high C-statistics but are based on simpler models with fewer degrees of freedom. Many of the models reporting high discrimination have been derived from and validated on lower-risk populations or within narrowly defined clinical cohorts.[6,7,9,16] For many of these models, 30-day mortality is measured postadmission and includes in-hospital mortality rather than being strictly postdischarge, an approach that boosts 30-day mortality C-statistics.[7] RSI models differ in being broadly applicable across the entire clinical populations, for both postdischarge 30-day mortality and in-hospital 30-day mortality.[8,10,11]

Demographic characteristics are an integral part of many risk adjustment models,[1–3] but as with RSI 1.0, demographic characteristics contribute almost nothing to discrimination of RSI 2.0 models. Furthermore, demographic characteristics alone were minimally predictive. Our results thus suggest that characteristics such as age are less important than patients' accumulated diagnostic codes. In other words, the diagnostic codes may well represent biologic age, which
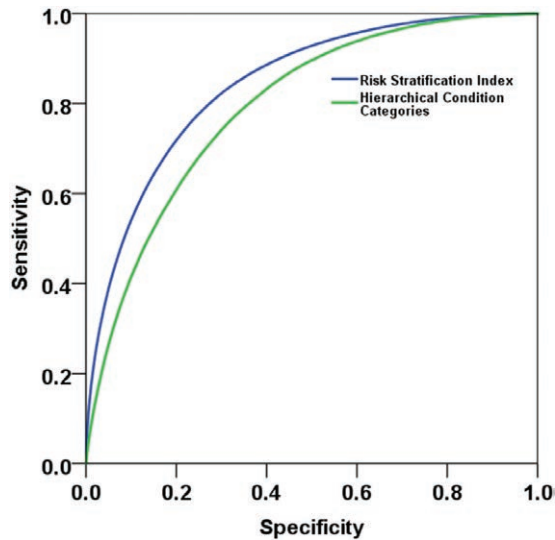
Fig. 3. Prediction of 1-yr mortality by Risk Stratification Index *versus* Hierarchical Condition Categories. The data are from a 5% sample of Medicare Analysis and Provider Review data set years 2012 to 2014.
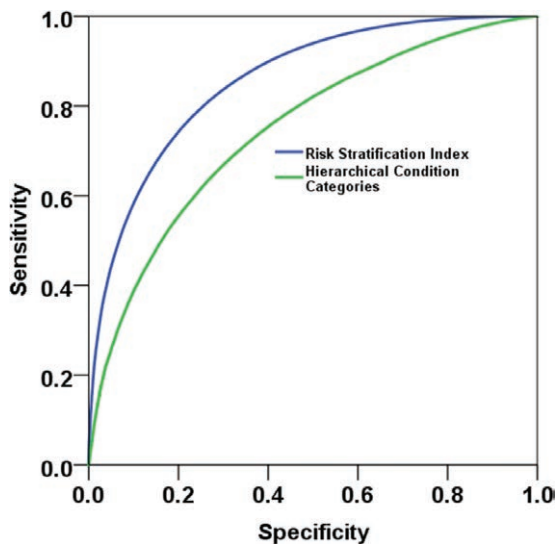

Fig. 5. Hierarchical Condition Categories (HCC) and Risk Stratification Index (RSI) Bootstrapping. The data are from an initial sample of 300,000 individuals from the Medicare Analysis and Provider Review File years 2007 to 2014. O/E = ratio of observed mortality against predicted.

Fig. 4. Prediction of median length-of-stay by Risk Stratification Index *versus* Hierarchical Condition Categories. The data are from a 5% sample of Medicare Analysis and Provider Review data set years 2012 to 2014.

Fig. 6. Hierarchical Condition Categories (HCC) 1-yr mortality calibration. HCC correlated with observed outcomes with an R^2 of 0.998 across 781 probability bins each of width 0.001. The data are from a Medicare Analysis and Provider Review data set years 2007 to 2014. Ninety percent of patients fall between predicted mortality of 0 and 55%, and 99% fall between 0 and 71%.

functions as a better predictor than chronologic age. Presumably any model that accurately assesses baseline health status will benefit little from including age, with the contribution of age being inversely related to strength of the model. In other words, the better a model estimates risk from other sources, the less age will contribute.

Despite overall similarities in patient characteristics, coding in 2012 to 2014 differed from other years—possibly because compliance with provisions of the Affordable Care Act enhanced coding. Interestingly, the years 2012 to 2014 also had the best discrimination for all HCC and RSI models. It thus seems likely that better coding resolution enhanced
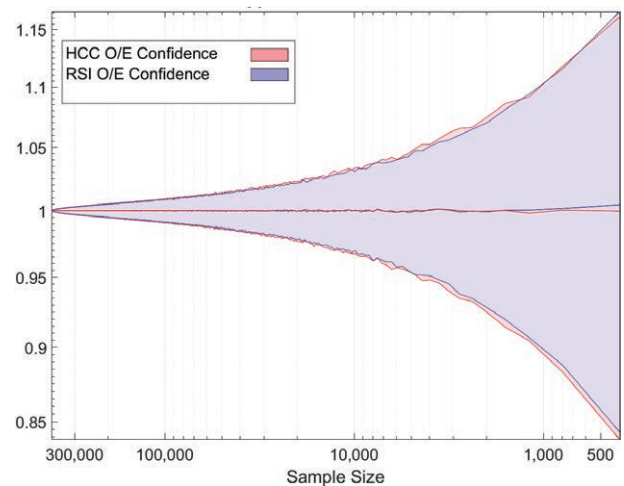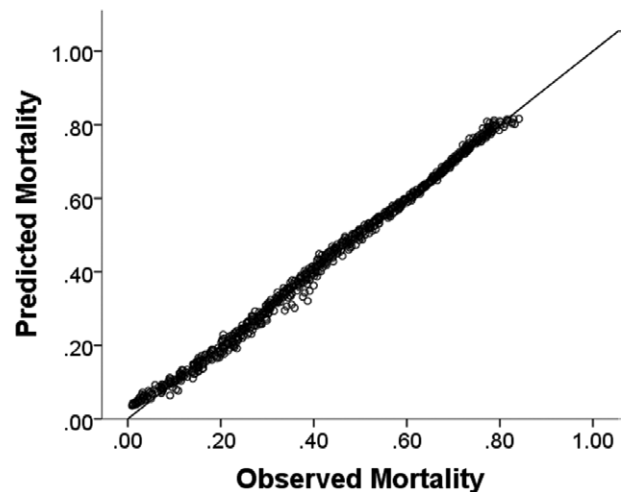
performance. Although RSI 2.0 has so far only been validated in patients at least 65 yr old, RSI 1.0 was highly predictive across a broad range of adult surgical patients,[8,10,11] suggesting RSI 2.0 models will perform well in younger populations.

In summary, updating RSI 1.0 models with more recent diagnostic codes and data improved performance only slightly, indicating that RSI models are remarkably stable across time. RSI 2.0 consistently discriminated better than HCC across all endpoints and was more precise after calibration. The consistently high performance and stability of RSI 2.0 models across multiple endpoints and time periods make
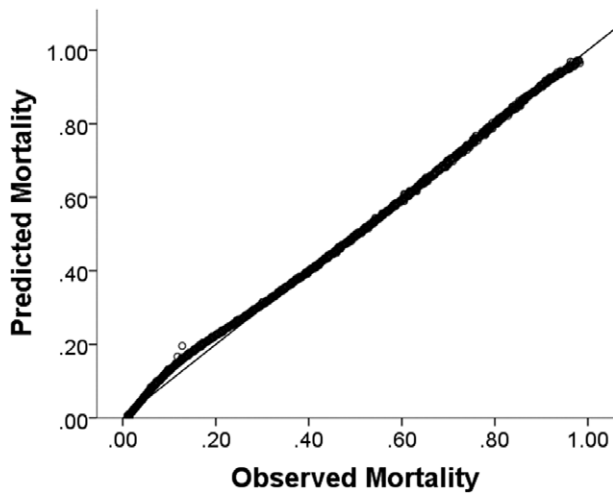
Fig. 7. Risk Stratification Index 1-yr calibration. One-year mortality correlated with observed outcomes with an R^2 of 0.998 across 974 probability bins each of width 0.001. The data are from a Medicare Analysis and Provider Review data set years 2007 to 2014. Ninety percent of patients fall between predicted mortality between 0 and 63%, and 99% fall between 0 and 89%.
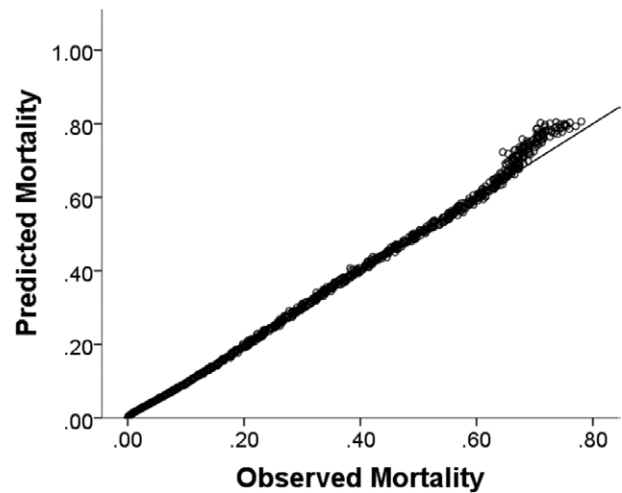

Fig. 9. Risk Stratification Index in-hospital mortality calibration. In-hospital correlated with observed outcomes with an R^2 of 0.999 across 806 probability bins each of width 0.001. The data are from a Medicare Analysis and Provider Review data set years 2007 to 2014. Ninety percent of patients fall between predicted mortality of 0 and 13%, and 99% fall between 0 and 61%.
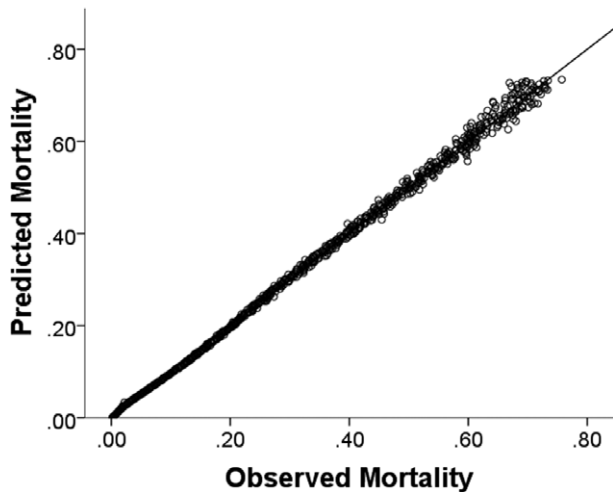

Fig. 8. Risk Stratification Index 30-day calibration. In-hospital correlated with observed outcomes with an R^2 of 0.999 across 734 probability bins each of width 0.001. The data are from a Medicare Analysis and Provider Review data set years 2007 to 2014. Ninety percent of patients fall between predicted mortality of 0 and 16%, and 99% fall between 0 and 51%.
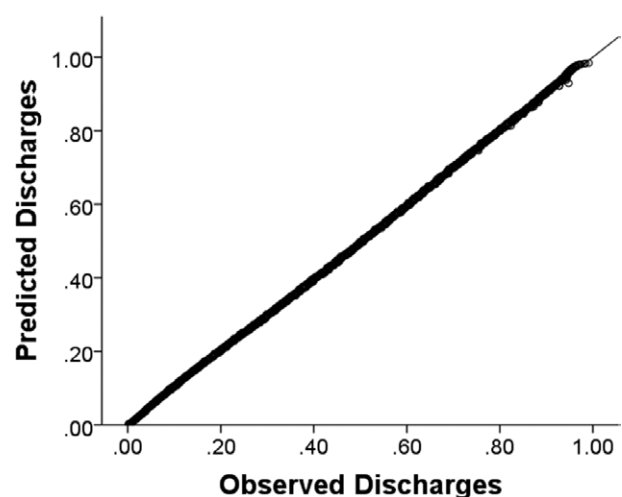

Fig. 10. Risk Stratification Index length-of-stay calibration. Length-of-stay correlated with observed outcomes with an R^2 of 1.000 across 983 probability bins each of width 0.001. The data are from a Medicare Analysis and Provider Review data set years 2007 to 2014. Ninety percent of the patients fall between predicted mortality of 0 and 91%, and 99% fall between 0 and 99%.

them attractive candidates for risk adjustment applications in health services research.

Research Support

Competing Interests

The authors declare no competing interests.

Correspondence

Address correspondence to Dr. Sessler: Cleveland Clinic, 9500 Euclid Ave., P77, Cleveland, Ohio 44195. DS@OR.org. Information on purchasing reprints may be found at www. anesthesiology.org or on the masthead page at the beginning of this issue. ANESTHESIOLOGY's articles are made freely accessible to all readers, for personal use only, 6 months from the cover date of the issue.

References

1. 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. Elixhauser A, Steiner C, Harris DR, Coffey RM: Comorbidity measures for use with administrative data. Med Care 1998; 36:8–27

3. Fitz-Henry J: The ASA classification and peri-operative risk. Ann R Coll Surg Engl 2011; 93:185–7

4. 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

5. 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

6. Le Manach Y, Collins G, Rodseth R, Le Bihan-Benjamin C, Biccard B, Riou B, Devereaux PJ, Landais P: Preoperative Score to Predict Postoperative Mortality (POSPOM): Derivation and validation. ANESTHESIOLOGY 2016; 124:570–9

7. Glance LG, Lustik SJ, Hannan EL, Osler TM, Mukamel DB, Qian F, Dick AW: The Surgical Mortality Probability Model: Derivation and validation of a simple risk prediction rule for noncardiac surgery. Ann Surg 2012; 255:696–702

8. 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

9. Dalton JE, Glance LG, Mascha EJ, Ehrlinger J, Chamoun N, Sessler DI: Impact of present-on-admission indicators on risk-adjusted hospital mortality measurement. ANESTHESIOLOGY 2013; 118:1298–306

10. Sigakis MJ, Bittner EA, Wanderer JP: 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. ANESTHESIOLOGY 2013; 119:525–40

11. Wahl KM, Moretti E, White W., Hale B, Gan T: Validation of a Risk-Stratification Index for Predicting 1-Year Mortality, Duke University Medical Center, Durham, North Carolina, United States, 2011

12. Chamoun GF, Li L, Chamoun NG, Saini V, Sessler DI: Validation and calibration of the Risk Stratification Index. ANESTHESIOLOGY 2017; 126:623–30

13. Services DoHaH: Hospital-acquired conditions and present on admission indicator reporting provision. Baltimore, Services CfMM, 2016

14. Gregory C, Pope M, John Kautter P, Melvin J, Ingber P, Sara Freeman M, Rishi Sekar B, Cordon Newhart M: Evaluation of the CMS-HCC Risk Adjustment Model. Baltimore, The Center for Medicare and Medicaid Services, 2011

15. Elixhauser A, Steiner C, Palmer L.: Clinical classifications software (CCS) 2015. Rockville, U.S. Agency for Healthcare Research and Quality, 2015

16. Kim M, Wall MM, Li G: Applying latent class analysis to risk stratification for perioperative mortality in patients undergoing intraabdominal general surgery. Anesth Analg 2016; 123:193–205