

OPEN

The Effect of Outcome Selection on the Performance of Prediction Models in Patients at Risk for Sepsis

Stephanie P. Taylor, MD, MS¹; Shih-Hsiung Chou, PhD²; Andrew D. McWilliams, MD, MPH³; Mark Russo, MD⁴; Alan C. Heffner, MD⁵; Stephanie Murphy, DO⁶; Susan L. Evans, MD, FACS, FCCM⁷; Whitney Rossman, MS²; Marc Kowalkowski, PhD²; on behalf of Acute Care Outcomes Research Network (ACORN) Investigators

Importance: Risk prediction models for patients with suspected sepsis have been derived on and applied to various outcomes, including readily available outcomes such as hospital mortality and ICU admission as well as longer-term mortality outcomes that may be more important to patients. It is unknown how selecting different outcomes influences model performance in patients at risk for sepsis.

¹Department of Internal Medicine, Atrium Health, Charlotte, NC.

²Center for Outcomes Research and Evaluation, Atrium Health, Charlotte, NC.

³Department of Internal Medicine, Division Hospital Medicine, Center for Outcomes Research and Evaluation, Atrium Health, Charlotte, NC.

⁴Department of Hepatology, Atrium Health, Charlotte, NC.

⁵Department of Internal Medicine, Emergency Medicine, Division of Critical Care, Atrium Health, Charlotte, NC.

⁶Department of Internal Medicine, Division of Hospital Medicine Transition Services, Atrium Health, Charlotte, NC.

⁷Department of Surgery, Atrium Health, Charlotte, NC.

Drs. Taylor and Chou contributed equally to the development of this article.

Supplemental digital content is available for this article. Direct URL citations appear in the HTML and PDF versions of this article on the journal's website (<http://journals.lww.com/ccejournal>).

Supported, in part, by grant from the Atrium Health Acute Care Outcomes Research Network. This evaluation is supported by internal funds from the Atrium Health Division of Adult Acute Care.

Dr. Murphy who is a speaker for Novartis regarding Sickle Cell Disease. The remaining authors have disclosed that they do not have any potential conflicts of interest.

For information regarding this article, E-mail: Shih-Hsiung.Chou@AtriumHealth.org

Copyright © 2020 The Authors. Published by Wolters Kluwer Health, Inc. on behalf of the Society of Critical Care Medicine. This is an open-access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

Crit Care Expl 2020; 2:e0078

DOI: 10.1097/CCE.000000000000078

Objectives: Evaluate the impact of outcome selection on risk model performance and weighting of individual predictor variables.

Design, Setting, and Participants: We retrospectively analyzed adults hospitalized with suspected infection from January 2014 to September 2017 at 12 hospitals.

Main Outcomes and Measures: We used routinely collected clinical data to derive logistic regression models for four outcomes: hospital mortality, composite ICU length of stay greater than 72 hours or hospital mortality, 30-day mortality, and 90-day mortality. We compared the performance of the models using area under the receiver operating characteristic curve and calibration plots.

Results: Among 52,184 admissions, 2,030 (4%) experienced hospital mortality, 6,659 (13%) experienced the composite of hospital mortality or ICU length of stay greater than 72 hours, 3,417 (7%) experienced 30-day mortality, and 5,655 (11%) experienced 90-day mortality. Area under the receiver operating characteristic curves decreased when hospital-based models were applied to predict 30-day (hospital mortality = 0.88–0.85; –0.03, composite ICU length of stay greater than 72 hours or hospital mortality = 0.90–0.81; –0.09) and 90-day mortality (hospital mortality = 0.88–0.81; –0.07, composite ICU length of stay greater than 72 hours or hospital mortality = 0.90–0.76; –0.14; all $p < 0.01$). Models were well calibrated for derived (root-mean-square error = 5–15) but not alternate outcomes (root-mean-square error = 8–35).

Conclusions and Relevance: Risk models trained to predict readily available hospital-based outcomes in suspected sepsis show poorer discrimination and calibration when applied to 30- and 90-day mortality. Interpretation and application of risk models for patients at risk of sepsis should consider these findings.

Key Words: calibration plot; infection; mortality; risk model; sepsis

Accurately predicting risk of poor outcomes for patients with suspected sepsis is critical to ensure high-risk patients receive appropriate aggressive therapy, timely allocation of intensive resources, and longer-term support.

Attempts to develop predictive tools that identify adverse outcomes have been limited by lackluster performance in heterogeneous settings. The variability in predictive accuracy of prediction tools may be partly due to different outcomes used in derivation and validation studies. For example, the original validation of the quick Sepsis-Associated Organ Failure Assessment (qSOFA) tool used hospital mortality as its primary outcome and a composite of hospital mortality and ICU stay greater than or equal to 3 days as a secondary outcome (1). Subsequent studies have used qSOFA to predict other outcomes, including acute organ dysfunction, ICU admission, hospital-, 28-day, 30-day, 90-day, 6-month, and 1-year mortality, or a composite outcome (2–4).

Different potential outcomes have advantages and disadvantages. Overall, mortality is an objective and patient-centered outcome. Hospital mortality is commonly chosen for risk prediction because it is the most reliably obtained mortality endpoint from hospital administrative data. However, the use of hospital mortality as an outcome may be subject to discharge bias that reflects discharge practices (e.g., to skilled nursing facilities, hospice) rather than true estimates of mortality (5, 6). Use of 30- or 90-day mortality can reduce discharge bias and have the advantage of reflecting consequences of the sepsis trajectory such as cognitive and functional decline (7), but these advantages are offset by the challenges healthcare systems face to reliably obtain outcomes data beyond the acute episode (e.g., data identification, linkage, and integration from disparate sources) (8). Finally, a common outcome used in predictive models for sepsis is ICU stay, which is readily measured (9). However, ICU admission is often discretionary and influenced by external factors such as hospital or provider practice patterns and bed availability (10, 11). In addition, using ICU admission as an outcome assumes that there are no unnecessary or false-positive ICU admission events. Finally, the advantage of a composite outcome including ICU admission and mortality is increased statistical power, but this outcome will likely be driven by more frequent ICU admission events. Although each of these outcomes have been used in different studies to assess sepsis risk models, it is not known whether the choice of outcome affects the performance of the models or the influence of predictor variables.

In this study, we derived predictive models for three mortality outcomes and one composite outcome of mortality and ICU admission and evaluated the impact of outcome selection on model performance and weighting of individual predictor variables. We hypothesized that predictors of these events will differ and that models derived for outcomes that are appealing to researchers (i.e., hospital mortality and the composite outcome of > 72-hr ICU stay and hospital mortality) will less accurately predict outcomes that are important to patients (i.e., 30- and 90-d mortality).

MATERIALS AND METHODS

Study Setting and Population

We conducted a retrospective cohort study by selecting adult (≥ 18 yr old) patients who presented to the emergency department (ED) and were hospitalized with clinically suspected infection between January 2014 and September 2017 at 12 acute care

hospitals within a large Southeast U.S. healthcare system. We adapted the definition of infection from the third international Consensus Definitions for Sepsis and Septic Shock (12), that is, an oral/parenteral antibiotic or bacterial culture ordered within 24 hours from ED admission and either: 1) a culture drawn first, antibiotics ordered within 48 hours or 2) antibiotics ordered first, culture ordered within 48 hours. We excluded hospital admissions for patients with antibiotics only ordered as preoperative infection prophylaxis, and patients with “do-not-resuscitate” or “do-not-intubate” orders within 24 hours from ED admission because of the higher potential for shifts in goals of care that might independently alter the risk of in-hospital mortality. The patient selection flow diagram is shown in **Figure 1**.

Data Collection

We extracted data from the healthcare system’s enterprise data warehouse, including sociodemographic and clinical characteristics (e.g., age, gender, race, insurance, diagnoses, prior healthcare utilization, clinical orders, laboratory values, and vital signs within the first 24hr from ED admission). We applied standard definitions to combine laboratory values and vital signs to generate qSOFA (1). We used *International Classification of Diseases* codes from healthcare encounters during the previous 12 months to categorize comorbidities and calculate weighted Charlson Comorbidity Index (CCI) scores (13).

Outcomes

We investigated four study outcomes: 1) hospital mortality, 2) a composite of greater than 72-hour ICU stay or hospital mortality, 3) 30-day mortality, and 4) 90-day mortality (**e-Fig. 1**, Supplemental Digital Content 1, <http://links.lww.com/CCX/A128>). Hospital mortality was determined using hospital administrative discharge disposition data and vital status records in the electronic health record (EHR). Mortality outcomes at 30 and 90 days were captured from the social security death index.

Statistical Analysis and Model Development

We used logistic regression to construct separate risk prediction models for each of the four outcomes adhering to the Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis statement on reporting predictive models (**e-Appendix 2**, Supplemental Digital Content 1, <http://links.lww.com/CCX/A128>) (14). All models were trained using hospital admissions from January 1, 2014, to December 31, 2016 (training dataset) and tested using hospital admissions between January 1, 2017, and September 30, 2017 (testing dataset). We included an initial set of 183 candidate variables. Prior to variable selection, laboratory values and vital signs captured as continuous variables were converted to categories based on quartiles and mean estimates in separate groups of hospital deaths and hospital survivors. Additionally, clinical thresholds for abnormal laboratory values (e.g., lactate > 2 mmol/L) and vital signs (e.g., respiratory rate > 20 breaths/min) were used to calculate the ratio of abnormal measurements to the total number of measurements for each variable within the first 24 hours of ED presentation. For selected continuous variables with less than 1% missing data, we also imputed

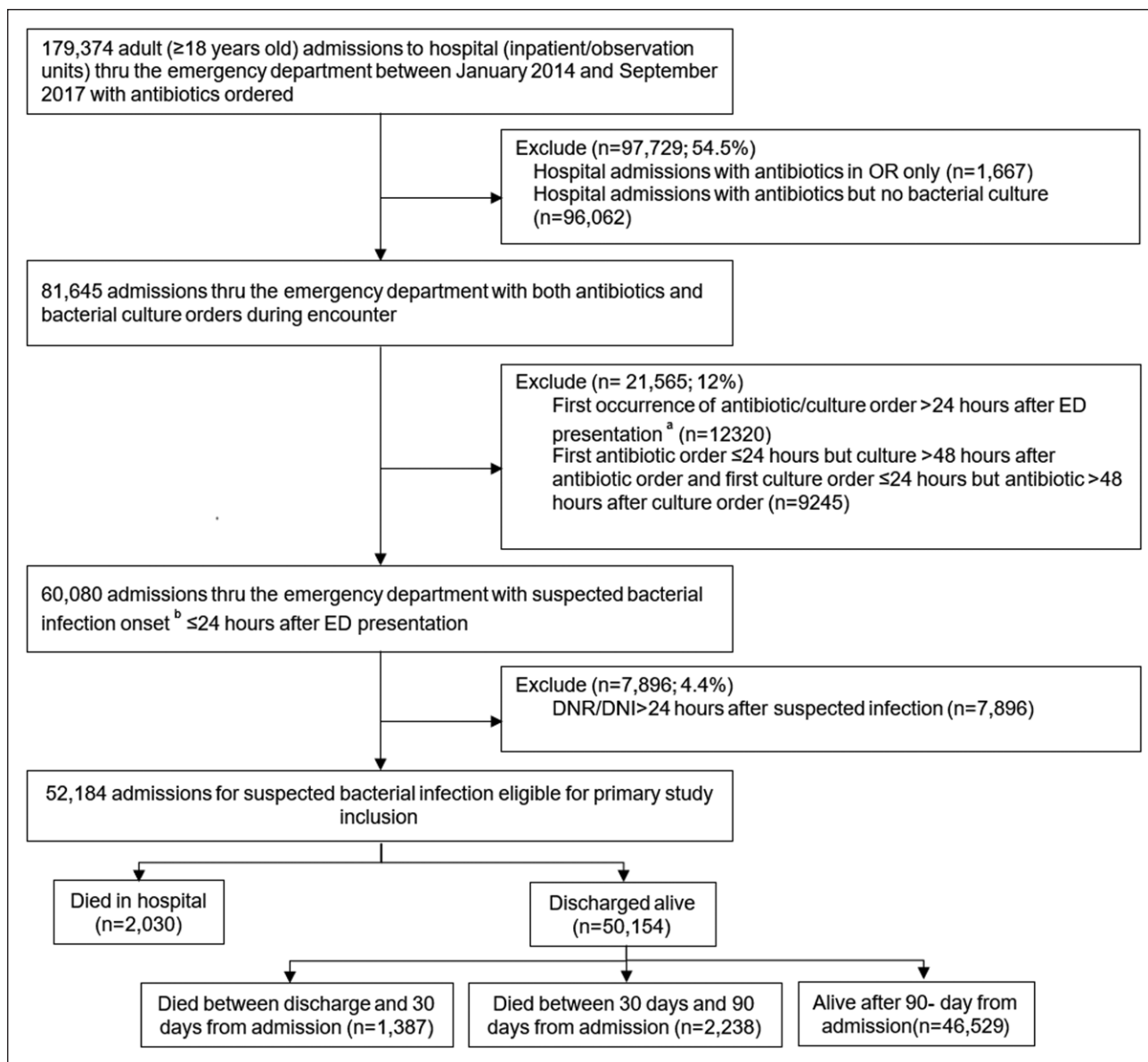


Figure 1. Selection of hospitalized patients with clinically suspected infection. There were 52,184 eligible admissions with suspected infection admitted through the emergency department (ED) to 12 study hospitals from January 2014 to September 2017. Suspected infection was defined by the following clinical criteria: oral/parenteral antibiotic or bacterial culture order within 24 hr of ED presentation and 1) culture drawn first, antibiotics ordered within 48 hr or 2) antibiotics ordered first, culture ordered within 48 hr. Patients with code status changes (i.e., orders placed for do not resuscitate [DNR] or do not intubate [DNI]) within 24 hr after ED presentation were excluded.

the mean value adjusted for age at admission, gender, race, CCI score, and body mass index. We tested the associations between variables with greater than 40% missing data and each outcome variable via chi-square tests. We then used a backward elimination approach and a prespecified threshold of p value of greater than 0.1 to exclude variables from the final risk models for each outcome. Additional details about model development, variable transformation and selection, and handling of missing values are included in **e-Appendix 1** and **e-Table 1** (Supplemental Digital Content 1, <http://links.lww.com/CCX/A128>).

We applied each risk prediction model to the four different mortality or composite outcomes derived from the training dataset and evaluated model performance on the independent testing dataset using discrimination statistics and calibration

plots. We measured the ability of each model to accurately differentiate between patients who did and did not have the indicated outcome using area under the receiver operating characteristic curve (AUC). All derived models were evaluated through k -fold cross-validation ($k = 10$) using the training dataset to resample and iteratively reestimate how accurately the prediction models perform using different, randomly assigned training and validation samples, a technique routinely applied to limit overfitting and selection bias (15). For fair comparison, the same random number seed was used when running the k -fold cross-validation. All models were applied to the testing dataset for final model performance assessment. We evaluated differences in model discrimination across outcomes by DeLong method (16).

We also generated calibration plots for both training and testing datasets to examine differences in the observed versus predicted event rates. The root-mean-square error (RMSE) was used to evaluate the differences between each of the derived models and the perfectly calibrated model. We compared observed 90-day mortality rates to predicted estimates from hospital mortality (died-hosp), composite ICU length of stay greater than 72 hours or hospital mortality (ICU72/died-hosp), and 30-day mortality (died30) models. All analyses were performed using SAS Enterprise Guide v7.1 (SAS Institute, Cary, NC) and R v3.5 (R Foundation for Statistical Computing, Vienna, Austria).

The study protocol was approved by the Atrium Health Institutional Review Board. A waiver of consent was granted based on minimal harm and general impracticability.

RESULTS

Table 1 shows demographic and clinical characteristics to illustrate the cohort. Among the 52,184 total admissions (41,856 unique patients) included in the study, 2,030 (4%) experienced hospital mortality, 6,659 (13%) experienced the composite of hospital mortality or ICU length of stay greater than 72 hours, 3,417 (7%) experienced 30-day mortality, and 5,655 (11%) experienced 90-day mortality. There were no statistically significant differences in the characteristics or outcomes for patients in the training ($n = 41,757$, 80%) and validation ($n = 10,427$, 20%) cohorts (all $p > 0.05$). Characteristics of patients experiencing each outcome and information on distribution of missingness are presented in e-Table 1 (Supplemental Digital Content 1, <http://links.lww.com/CCX/A128>) and the variables selected for each model by backward elimination are shown in e-Table 2 (Supplemental Digital Content 1, <http://links.lww.com/CCX/A128>).

Model discrimination estimates for the four models are shown in **Table 2** and e-Table 3 (Supplemental Digital Content 1, <http://links.lww.com/CCX/A128>). In general, model discrimination was highest for predicting the composite ICU72/died-hosp (AUCs = 0.86–0.90), followed by died-hosp (AUC = 0.88), 30-day mortality (AUCs = 0.81–0.87), and 90-day mortality (AUCs = 0.76–0.85) in the testing dataset. Model discrimination decreased significantly when hospital-based outcome models (died-hosp, ICU72/died-hosp) were applied to predict 30- and 90-day mortality ($p < 0.01$). We observed the largest decline in performance when the ICU72/died-hosp model was used to predict 90-day mortality (-0.14 ; $p < 0.01$).

Calibration plots comparing observed versus predicted risks of each outcome are shown in **Figure 2** and e-**Figure 1** (Supplemental Digital Content 1, <http://links.lww.com/CCX/A128>). The identity line is indicative of a perfect model, in which the observed number of events is equal to the predicted risk across the range of estimated values (i.e., 0–100). Visual inspection of observed-to-expected (OE) risk estimates indicates that all risk models were well calibrated for their own outcomes (RMSE = 5–9) except died-hosp model (RMSE = 15) on the testing dataset. However, models were miscalibrated when predicting other outcomes (RMSE = 8–35). For example, the composite ICU72/died-hosp model overpredicted 90-day mortality (OE risk: 18% vs 25%). Conversely, the died-hosp (OE

risk: 42% vs 25%) and died30 (OE risk: 37% vs 25%) models underpredicted 90-day mortality.

DISCUSSION

In this study of patients at risk for sepsis, we illustrate several considerations for applying predictive models to outcomes other than those from which the models were initially derived. As hypothesized, models derived using readily available outcomes (i.e., hospital mortality and/or 72-hr ICU stay) showed incremental decrease in discrimination when applied to more patient-centered outcomes (i.e., 30- and 90-d mortality), although the absolute difference in AUC was small and of unknown clinical significance.

Second, our results highlight the potential for miscalibration when applying models to alternate outcomes. Often model evaluations focus on measures of discrimination (i.e., Do patients with the outcome have higher risk predictions than those without?) as measured by concordance statistics (e.g., AUC) more than calibration (i.e., Do x of 100 patients, with a risk prediction of $x\%$ experience the outcome?), which is assessed graphically and with OE ratios. A model with poor calibration could have important implications when applied in clinical situations. For example, the ICU 72/died-hosp model overpredicted 30-day mortality, which could lead to unnecessarily providing additional resources to patients deemed high risk or inappropriately counseling to deescalate care. Conversely, the died-hosp model underpredicted 30- and 90-day mortality. Thus, models derived using this outcome might not accurately identify high-risk patients who survive hospitalization but ultimately succumb to sepsis following longer-term sequelae. We note that in certain situations, clinical care has been improved by simply categorizing patients into broad risk categories (e.g., low, medium, and high), which would diminish the importance of precision in calibration (17). As such, the decision to apply risk models to outcomes other than what was studied will be nuanced.

Our study has several key strengths. First, it included over 50,000 admissions and a heterogeneous patient population. Second, our study investigated predictors obtained within the first 24 hours of presentation, integrating both clinical and administrative data in contrast to other studies that include only physiologic variables or only variables available too late for an intervention to alter the clinical trajectory. This near-real-time risk modeling strategy can optimize value by matching resources to high-risk patients early in the hospitalization. It can also be valuable to inform patient selection for future pragmatic trials. This is the first study to our knowledge to investigate the differences in predictors of hospital mortality, 30- and 90-day mortality, and composite hospital mortality and ICU length of stay and the implications the selected outcome has on model performance. This study uniquely evaluates model calibration across outcomes, an important model parameter that is frequently overlooked.

There are important limitations to this study. Notably, our study was conducted within one large integrated healthcare system, which may limit external generalizability. However, the population was selected from 12 hospitals with diverse characteristics. Our cohort may reflect an overall less severely ill population due to our selection strategy. We deliberately applied broad clinical data to

TABLE 1. Characteristics of Clinically Suspected Infection Study Cohort

Characteristics	Overall	Training Dataset (January 1, 2014~December 31, 2016)	Testing Dataset (January 1, 2017~September 30, 2017)
<i>n</i>	52,184	41,757	10,427
Age at admission, median (IQR)	60 (46–71)	60 (46–72)	60 (46–71)
Male gender, <i>n</i> (%)	23,868 (45.7)	18,906 (45.3)	4,962 (47.6)
Race, <i>n</i> (%)			
Black	13,485 (25.8)	10,778 (25.8)	2,707 (26.0)
White	35,661 (68.3)	28,639 (68.6)	7,022 (67.3)
Other	3,038 (5.8)	2,340 (5.6)	698 (6.7)
Insurance, <i>n</i> (%)			
Medicaid	8,820 (16.9)	7,348 (17.6)	1,472 (14.1)
Medicare	27,125 (52.0)	21,645 (51.8)	5,480 (52.6)
Commercial	11,242 (21.5)	8,938 (21.4)	2,304 (22.1)
Self-pay	4,331 (8.3)	3,218 (7.7)	1,113 (10.7)
Other/unknown	666 (1.3)	608 (1.5)	58 (0.6)
CCI, median (IQR)	4 (2–7)	4 (2–7)	4 (2–7)
CCI ≥ 5, <i>n</i> (%)	25,262 (48.4)	20,130 (48.2)	5,132 (49.2)
qSOFA ^a , median (IQR)	1 (0–1)	1 (0–1)	1 (0–1)
qSOFA ≥ 2, <i>n</i> (%)	9,829 (18.8)	7,756 (18.6)	2,073 (19.9)
GCS ^a , median (IQR)	15 (14–15)	15 (14–15)	15 (14–15)
GCS < 15, <i>n</i> (%)	15,008 (28.8)	11,822 (28.3)	3,186 (30.6)
SBP ^a , median (IQR)	102 (90–115)	102 (90–115)	102 (90–115)
SBP > 90 mm Hg, <i>n</i> (%)	39,058 (74.8)	31,283 (74.9)	7,775 (75.6)
Respiratory rate ^a , median (IQR)	23 (20–28)	23 (20–28)	23 (20–29)
Respiratory rate > 20 breaths/min, <i>n</i> (%)	32,629 (62.5)	26,106 (62.5)	6,523 (62.6)
Lactate ^a , median (IQR)	1.7 (1.2–2.8)	1.7 (1.2–2.8)	1.7 (1.1–2.8)
Lactate > 2 mmol/L, <i>n</i> (%)	13,370 (25.6)	10,714 (25.7)	2,656 (25.5)
Clinical outcomes, <i>n</i> (%)			
ICU admitted	15,083 (28.9)	12,117 (29.0)	2,966 (28.4)
ICU LOS > 72 hr	5,716 (11.0)	4,584 (11.0)	1,132 (10.9)
Hospital mortality	2,030 (3.9)	1,638 (3.9)	392 (3.6)
Hospital mortality or ICU LOS > 72 hr	6,659 (12.8)	5,347 (12.8)	1,312 (12.6)
30-d mortality from admission	3,417 (6.5)	2,756 (6.6)	661 (6.3)
90-d mortality from admission	5,655 (10.8)	4,532 (10.9)	1,123 (10.8)

CCI = Charlson Comorbidity Index Score, GCS = Glasgow Coma Scale, IQR = interquartile range, LOS = length of stay, qSOFA = quick Sepsis-Associated Organ Failure Assessment, SBP = systolic blood pressure.

^aPhysiologic measures and laboratory values captured within 24 hr from emergency department presentation.

define a population of patients with suspected infection to mirror data that would be available at the time of clinical decision-making. Despite this, observed mortality rates were similar to the population used to originally derive and validate the qSOFA tool (1). Further, although we evaluated risk models developed from

routinely collected data elements that are available in the EHR, the complexity of the models derived to predict the study outcomes requires computational bandwidth and may not be readily recreated in other settings. Finally, we used the Social Security Death Master File to track 30- and 90-day mortality, which may underestimate

TABLE 2. Model Performance Across Different Selection of Outcomes on Testing Dataset

Model	Outcome			
	Died-Hosp	ICU72/Died-Hosp	Died30	Died90
Died-hosp	0.88 (0.87–0.90)	0.86 (0.85–0.87)	0.85 (0.84–0.87)	0.81 (0.80–0.83)
ICU72/died-hosp	–	0.90 (0.90–0.91)	0.81 (0.80–0.83)	0.76 (0.75–0.77)
Died30	–	–	0.87 (0.85–0.88)	0.84 (0.83–0.85)
Died90	–	–	–	0.85 (0.84–0.86)

Died30 = 30-d mortality, died90 = 90-d mortality, died-hosp = hospital mortality, ICU72/died-hosp = composite of > 72-hr ICU stay or hospital mortality. Model discrimination presented as area under receiver operating characteristic curve (AUC) and 95% CIs.

AUC and 95% CI generated by DeLong method.

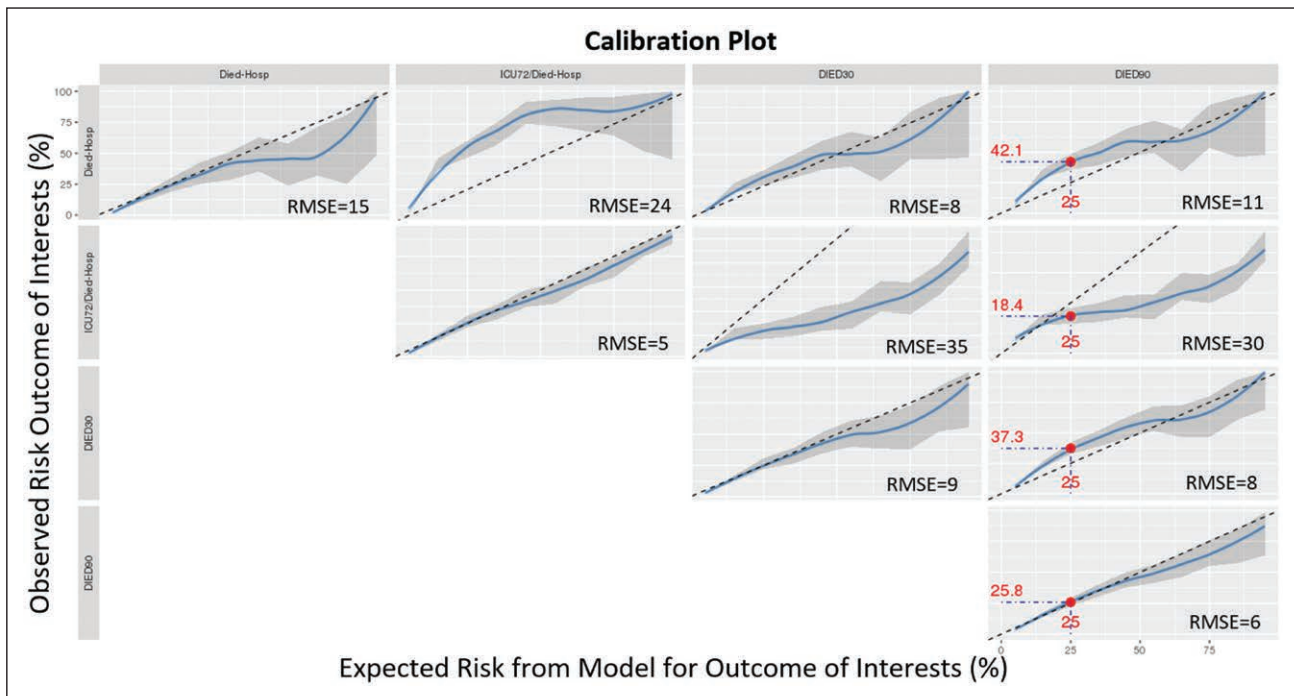


Figure 2. Calibration plots for died-hosp, ICU72/died-hosp, died30, and died90 model against outcomes of interest on testing dataset. Calibration plots are depicted for each model and outcome pair. The *x*-axis of all inner plots is the expected risk (%) for each of the outcomes of interest, whereas the *y*-axis represents observed risk (%) for each of the outcomes. The *identity line* is indicated by a *dashed line* and represents a perfectly calibrated model, in which the observed number of events are equal to the predicted number of events. The *solid line* indicates the actual number of observed events across the range of predicted risk values (i.e., 0–100). The area within the 95% confidence band around each of the observed estimates is *shaded gray*. Root-mean-square error (RMSE) between prediction models (*solid line*) and the perfectly calibrated model (*dashed line*). The *circles* illustrate examples comparing observed 90-d mortality versus expected risk predicted by 1) died-hosp, 2) ICU72/died-hosp, and 3) died30 models. At expected risks of 25%, died-hosp and died30 models underpredicted the 90-d mortality risk (observed risk = 42% and 37%, respectively), while the ICU72/died-hosp model overpredicted the 90-d mortality risk (observed risk = 18%). Died30 = 30-d mortality, died90 = 90-d mortality, died-hosp = hospital mortality, ICU72/died-hosp = composite of > 72-hr ICU stay or hospital mortality.

mortality in some populations (18). Our hybrid approach that combines internal health system and national mortality data attempts to overcome previously described limitations, but it is possible that our data still underestimate the different mortality rates.

CONCLUSIONS

Previous studies on risk stratification of suspected sepsis patients have developed prediction models using different outcomes. This variability in the outcome of interest, along with differences in patient populations, definitions of model performance, and study design makes it challenging to compare model performance across

studies. Our work provides clarity by demonstrating how model performance and predictors can differ depending on the outcome studied. We illustrate the trade-off in using models built on readily available hospital outcomes data to predict longer-term events that may be more important to patients. Clinical application of sepsis risk models and future studies should consider these findings.

ACKNOWLEDGMENTS

We acknowledge the collaboration of the Atrium Health Acute Care Outcomes Research Network Investigators listed here (in alphabetical order): Ryan Brown, MD; Larry Burke, MD; Shih-Hsiung

Chou, PhD; Kyle Cunningham, MD; Susan L. Evans, MD; Scott Furney, MD; Michael Gibbs, MD; Alan Heffner, MD; Timothy Hetherington, MS; Daniel Howard, MD; Marc Kowalkowski, PhD; Scott Lindblom, MD; Andrea McCall; Lewis McCurdy, MD; Andrew McWilliams, MD, MPH; Stephanie Murphy, DO; Alfred Papali, MD; Christopher Polk, MD; Whitney Rossman, MS; Michael Runyon, MD; Mark Russo, MD; Melanie Spencer, PhD; Brice Taylor, MD; Stephanie Taylor, MD, MS.

REFERENCES

1. Seymour CW, Liu VX, Iwashyna TJ, et al: Assessment of clinical criteria for sepsis: For the third international consensus definitions for sepsis and septic shock (sepsis-3). *JAMA* 2016; 315:762–774
2. Hwang SY, Jo IJ, Lee SU, et al: Low accuracy of positive qsofa criteria for predicting 28-day mortality in critically ill septic patients during the early period after emergency department presentation. *Ann Emerg Med* 2018; 71:1–9.e2
3. Raith EP, Udy AA, Bailey M, et al; Australian and New Zealand Intensive Care Society (ANZICS) Centre for Outcomes and Resource Evaluation (CORE): Prognostic accuracy of the SOFA score, SIRS criteria, and qSOFA score for in-hospital mortality among adults with suspected infection admitted to the intensive care unit. *JAMA* 2017; 317:290–300
4. Churpek MM, Snyder A, Sokol S, et al: Investigating the impact of different suspicion of infection criteria on the accuracy of quick sepsis-related organ failure assessment, systemic inflammatory response syndrome, and early warning scores. *Crit Care Med* 2017; 45:1805–1812
5. Carey JS, Parker JP, Robertson JM, et al: Hospital discharge to other healthcare facilities: Impact on in-hospital mortality. *J Am Coll Surg* 2003; 197:806–812
6. Vasilevskis EE, Kuzniewicz MW, Dean ML, et al: Relationship between discharge practices and intensive care unit in-hospital mortality performance: Evidence of a discharge bias. *Med Care* 2009; 47:803–812
7. Prescott HC, Costa DK: Improving long-term outcomes after sepsis. *Crit Care Clin* 2018; 34:175–188
8. Chang H: Making sense of the big picture: Data linkage and integration in the era of big data. *Healthc Inform Res* 2018; 24:251–252
9. Song JU, Sin CK, Park HK, et al: Performance of the quick sequential (sepsis-related) organ failure assessment score as a prognostic tool in infected patients outside the intensive care unit: A systematic review and meta-analysis. *Crit Care* 2018; 22:28
10. Town JA, Churpek MM, Yuen TC, et al: Relationship between ICU bed availability, ICU readmission, and cardiac arrest in the general wards. *Crit Care Med* 2014; 42:2037–2041
11. Robert R, Coudroy R, Ragot S, et al: Influence of ICU-bed availability on ICU admission decisions. *Ann Intensive Care* 2015; 5:55
12. Singer M, Deutschman CS, Seymour CW, et al: The third international consensus definitions for sepsis and septic shock (sepsis-3). *JAMA* 2016; 315:801–810
13. Charlson ME, Pompei P, Ales KL, et al: A new method of classifying prognostic comorbidity in longitudinal studies: Development and validation. *J Chronic Dis* 1987; 40:373–383
14. Collins GS, Reitsma JB, Altman DG, et al: Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD). *Ann Intern Med* 2015; 162:735–736
15. Steyerberg EW: *Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating*. New York, Springer Science & Business Media, 2008
16. DeLong ER, DeLong DM, Clarke-Pearson DL: Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach. *Biometrics* 1988; 44:837–845
17. Spertus JA, Decker C, Gialde E, et al: Precision medicine to improve use of bleeding avoidance strategies and reduce bleeding in patients undergoing percutaneous coronary intervention: Prospective cohort study before and after implementation of personalized bleeding risks. *BMJ* 2015; 350:h1302
18. Levin MA, Lin HM, Prabhakar G, et al: Alive or dead: Validity of the social security administration death master file after 2011. *Health Serv Res* 2019; 54:24–33