External validation of the COVID-19 4C mortality score in an urban United States cohort

Joshua M. Riley, BS[1], Patrick J. Moeller, MPH[2], Albert G. Crawford, PhD[2], Joseph W. Schaefer, BChE[1], Dianna R. Cheney-Peters, MD[3], Chantel M. Venkataraman, MD[4], Chris J. Li, BS[1], Christa M. Smaltz, MD[4], Conor G. Bradley, BS[1], Crystal Y. Lee, MPH[1], Danielle M. Fitzpatrick, MD[4], David B. Ney, BA[1], Dina S. Zaret, MD[1], Divya M. Chalikonda, MD[4], Joshua D. Mairose, BS[1], Kashyap Chauhan, MD[4], Margaret V. Szot, MD[4], Robert B. Jones, MD[4], Rukaiya Bashir-Hamidu, MD[4], Shuji Mitsuhashi, MD[4] and Alan A. Kubey, MD[3,5]

[1] Sidney Kimmel Medical College, Thomas Jefferson University, Philadelphia, PA, USA; [2] Jefferson College of Population Health, Thomas Jefferson University, Philadelphia, PA, USA; [3] Division of Hospital Medicine, Department of Medicine, Thomas Jefferson University Hospital, Philadelphia, PA, USA; [4] Internal Medicine Residency, Department of Medicine, Thomas Jefferson University Hospital, Philadelphia, PA, USA; [5] Division of Hospital Internal Medicine, Department of Internal Medicine, Rochester, MN, USA

ABSTRACT

Background: Identifying patients at risk for mortality from COVID-19 is crucial to triage, clinical decision-making, and the allocation of scarce hospital resources. The 4C Mortality Score effectively predicts COVID-19 mortality, but it has not been validated in a United States (U.S.) population. The purpose of this study is to determine whether the 4C Mortality Score accurately predicts COVID-19 mortality in an urban U.S. adult inpatient population.

Methods: This retrospective cohort study included adult patients admitted to a single-center, tertiary care hospital (Philadelphia, PA) with a positive SARS-CoV-2 PCR from 3/01/2020 to 6/06/2020. Variables were extracted through a combination of automated export and manual chart review. The outcome of interest was mortality during hospital admission or within 30 days of discharge.

Results: This study included 426 patients; mean age was 64.4 years, 43.4% were female, and 54.5% self-identified as Black or African American. All-cause mortality was observed in 71 patients (16.7%). The area under the receiver operator characteristic curve of the 4C Mortality Score was 0.85 (95% confidence interval, 0.79-0.89).

Conclusions: Clinicians may use the 4C Mortality Score in an urban, majority Black, U.S. inpatient population. The derivation and validation cohorts were treated in the pre-vaccine era so the 4C Score may over-predict mortality in current patient populations. With stubbornly high inpatient mortality rates, however, the 4C Score remains one of the best tools available to date to inform thoughtful triage and treatment allocation.

Key Indexing Terms: COVID-19; Evidence-based medicine; Risk; Triage. **[Am J Med Sci 2022;364(4):409–413.]**

INTRODUCTION

The management of patients with COVID-19 is challenging for front-line healthcare providers given limited validated, evidence-based clinical decision support. Determining patient mortality risk is critical for effective triage, management, and discharge decision making. Numerous COVID-19 risk prediction models of varying country of origin, population demographic, and robustness exist.[1-3] The 4C Mortality Score, created by the International Severe Acute Respiratory and emerging Infections Consortium (ISARIC), is one of the largest-scale, high-performing predictive models published to date.[4]

The initial 4C Mortality Score study included adult patients hospitalized with COVID-19 across 260 hospitals in England, Wales, and Scotland, utilizing eight weighted variables – age, sex, select comorbidities, respiratory rate, peripheral oxygen saturation, Glasgow Coma Scale (GCS), blood urea nitrogen (BUN), and C-reactive protein (CRP) – to predict 30-day inpatient

mortality. The area under receiver operator characteristic curve (AUC) of the derivation (n=35,463) and validation (n=22,361) cohorts was 0.79 (95% confidence interval [CI], 0.78-0.79) and 0.77 (95% CI, 0.76-0.77), respectively. This model outperformed 16 risk stratification scores, including SOFA, NEWS, CURB-65 and four novel COVID-19 risk scores.[4]

The 4C Mortality Score has since been validated in a London cohort (n=8239; AUC=0.68 [95% CI, 0.67-0.69])[5]; a Dutch emergency department cohort (n=403; AUC=0.84 [95% CI, 0.79-0.88])[6]; a Brazilian and Spanish cohort (n=1363; AUC = 0.78 [95% CI, 0.75-0.81])[7]; and an Italian cohort of greater than 60 years old (n=210, AUC=0.80 [95% CI, 0.74-0.85]).[8] For United States clinicians to confidently utilize the 4C Mortality Score, however, the model must be validated in a population similar to its intended implementation, particularly in areas with minorities under-represented in prior research. The objective of this study is to determine whether the 4C Mortality Score is a clinically appropriate model to predict hospital mortality in United States patients with COVID-19.

METHODS

Study population

Patients included in this study had a positive SARS-CoV-2 PCR test result and were admitted through the emergency department to Thomas Jefferson University Hospital (TJUH), a tertiary care academic medical center in Philadelphia, PA, between March 1st and June 6th, 2020. Symptomatic and asymptomatic patients were included. Patients transferred from outside hospitals were excluded. Patients under the age of 18, pregnant, and/or incarcerated were excluded from this study. This study was approved by the TJUH Institutional Review Board (IRB#: 20E.737).

Variables and outcome

The following 4C Mortality Score comorbidities were extracted from electronic health records (EHR) through manual chart review: chronic cardiac disease, chronic respiratory disease (excluding asthma), chronic kidney disease (estimated glomerular filtration rate $\leq$ 30 mL/min/1.73m^2), liver disease, connective tissue disease, diabetes mellitus, HIV/AIDS, active malignancy, and obesity (body mass index [BMI] $\geq$ 30 kg/m^2). These data were obtained from the last discharge summary, the emergency department note, the admission history and physical, the EHR "Past Medical History" and/or "Active Problems" sections. All comorbidities were retrieved by two separate researchers; a third independent reviewer adjudicated data discrepancies. The study database was created prior to the release of the 4C Mortality Score; therefore, two comorbidities – dementia and "chronic neurological disease " – were not included in this study. Respiratory rate, peripheral oxygen saturation, GCS, BUN, and CRP, within 24 hours of initial hospital encounter were digitally exported. The initial hospital encounter was defined as the first time-value of vitals entered in the EHR for the admission of interest. The primary outcome was a composite outcome of mortality during admission or mortality within 30 days of discharge. Mortality status was determined by chart review which included deaths occurring at other hospitals using the same EHR (the three largest academic medical centers in Philadelphia are included). This study is reported consistently with the Transparent Reporting of a multivariate prediction model for Individual Prognosis or Diagnosis guidelines.

Statistical analysis

Data was missing for less than 1.4% of variables. Values were missing for four variables of the 4C Score: CRP, GCS, BUN, and obesity (Table 1). Missing data was assumed to be missing at random. Multiple imputation methods were applied to the 4C Score components using R (Version 4.0.4, R Project for Statistical Computing) and the Multivariate Imputation by Chained Equations package (MICE, v3.13.0). Twenty imputed datasets were created, using predictive mean matching for numerical variables (e.g., CRP) and logistic regression for binary variables (e.g., obesity).

Adjusted 4C Mortality Scores were computed for each record using the data in each imputed dataset. Mortality probability for each 4C Score was sourced from the ISARIC online 4C calculator (https://isaric4c.net/risk/v2/). Receiver operating characteristic curves were calculated for each imputed dataset comparing the predicted 4C Score and actual mortality for each record. Area under the receiver operator curve and Brier score, another measure of goodness of fit (ranging from 0 to 1; smaller values indicate superior calibration), were calculated for each curve. AUC, corresponding 95% CI, and Brier score were pooled for all 20 imputation results using Rubin's rule.[9]

RESULTS

Population demographics and outcomes

There were 426 patients included in this study. Mean age was 64.4 years; 43.4% were female; 54.5% were Black or African American (hereafter "Black"). Cohort demographics, clinical variables, and outcomes are presented in Table 1. The primary outcome of mortality during admission or within 30 days from discharge occurred in 16.7% of patients (71/426 patients). Of the 71 deaths, 10 occurred after discharge. Of deaths occurring after the index admission, the average time from index admission to death was 13.9 days (minimum: 7 days; maximum 21 days). The incidence of other measures of clinical interest (i.e., ICU status, discharge to hospice) can also be seen in Table 1.

TABLE 1. Demographics, outcomes, and clinical variables of the study cohort and 4C Mortality Score cohort.

Characteristic	External Validation Cohort		4C Score Derivation Cohort
	Number of patients (% of total)	% Missing	Number of patients (% of total)
Age	-	0	-
<50	91 (21.4)	-	4876 (13.8)
50-69	168 (39.4)	-	10183 (28.9)
70-79	68 (16.0)	-	8017 (22.7)
≥80	99 (23.2)	-	12 201 (34.6)
Sex at birth	-	0	-
Female	184 (43.4)	-	14741 (41.7)
Male	241 (56.6)	-	20615 (58.3)
Ethnicity	-	0	-
Black	232 (54.5)	-	1256 (3.9)
White	109 (25.6)	-	26300 (82.2)
Other ethnicity	85 (19.9)	-	2513 (7.9)
Outcomes			
Mortality	71 (16.7)	0	11426 (32.2)
ICU Status	85 (19.9)	-	-
Discharge to Hospice	14 (3.3)	-	-
Comorbities	**No. of pts. (%)**	**% Missing**	**No. of pts. (%)**
Chronic cardiac disease	123 (28.8)	0	10513 (31.8)
Chronic respiratory disease	75 (17.6)	0	5830 (17.7)
Chronic kidney disease	41 (9.6)	0	5653 (17.2)
Active cancer	28 (6.6)	0	3312 (10.2)
Liver disease	17 (4.0)	0	604 (1.9)
Obesity	172 (40.4)	1.2	3414 (11.4)
Diabetes	162 (38.0)	0	8487 (26.0)
Number of Comorbities			
0	94 (22.1)	-	8497 (24.0)
1	142 (33.3)	-	9941 (28.0)
≥2	190 (44.6)	-	17025 (48.0)
Continuous variables	**Median (IQR)**	**% Missing**	**Median (IQR)**
Respiratory rate (breaths/min)	20 (5)	0	22 (9)
Oxygen saturation (%)	96 (5)	0	94 (6)
Glasgow Coma Scale	15 (0)	7.7	15 (0)
Blood urea nitrogen (mg/dL)	18.0 (19.0)	1.4	19.6 (17.6)
C-reactive protein (mg/dL)	5.9 (8.2)	11.7	8.5 (12.2)

Performance of 4C mortality score

The AUC for the pooled dataset was 0.85 (95% CI, 0.79-0.89). The Brier score for the dataset was 0.246. The calibration curve for all imputed datasets is plotted in Fig. 1; calibration curves for each individual imputation are plotted in the Supplemental Materials.

DISCUSSION

With an AUC of 0.85 (95% CI, 0.79-0.89) and Brier score of 0.246, the 4C Mortality Score performed well in a United States urban population with a majority Black patient cohort. These findings are consistent with the results of external validation studies in other countries.[6-8,10,11] We note a non-statistically significant trend toward overprediction of mortality, particularly visually noticeable in 4C scores <15 as seen in Fig. 1. This may be non-significant data variation, but suggests the need for further research in larger United States populations to investigate. We suspect that the 4C Score will tend to over-predict mortality in current populations given the increased availability of vaccines and effective outpatient and inpatient treatments relative to the study population (prior to June 7, 2020).[12-20] Given stubbornly high inpatient mortality rates to date, however, this assumption may not hold true.

Notably, this dataset utilized manual chart review, which was shown to outperform automated EHR data export for the sensitivity of capturing comorbidities within this study population.[21] This may have led to higher scores in our validation cohort relative to the original derivation cohort or other validation cohorts. Further strengths of this study include the representation of the racial diversity of the center's patient population, with a majority of the study cohort identifying as Black (54.5%). Recent research into the association of race,

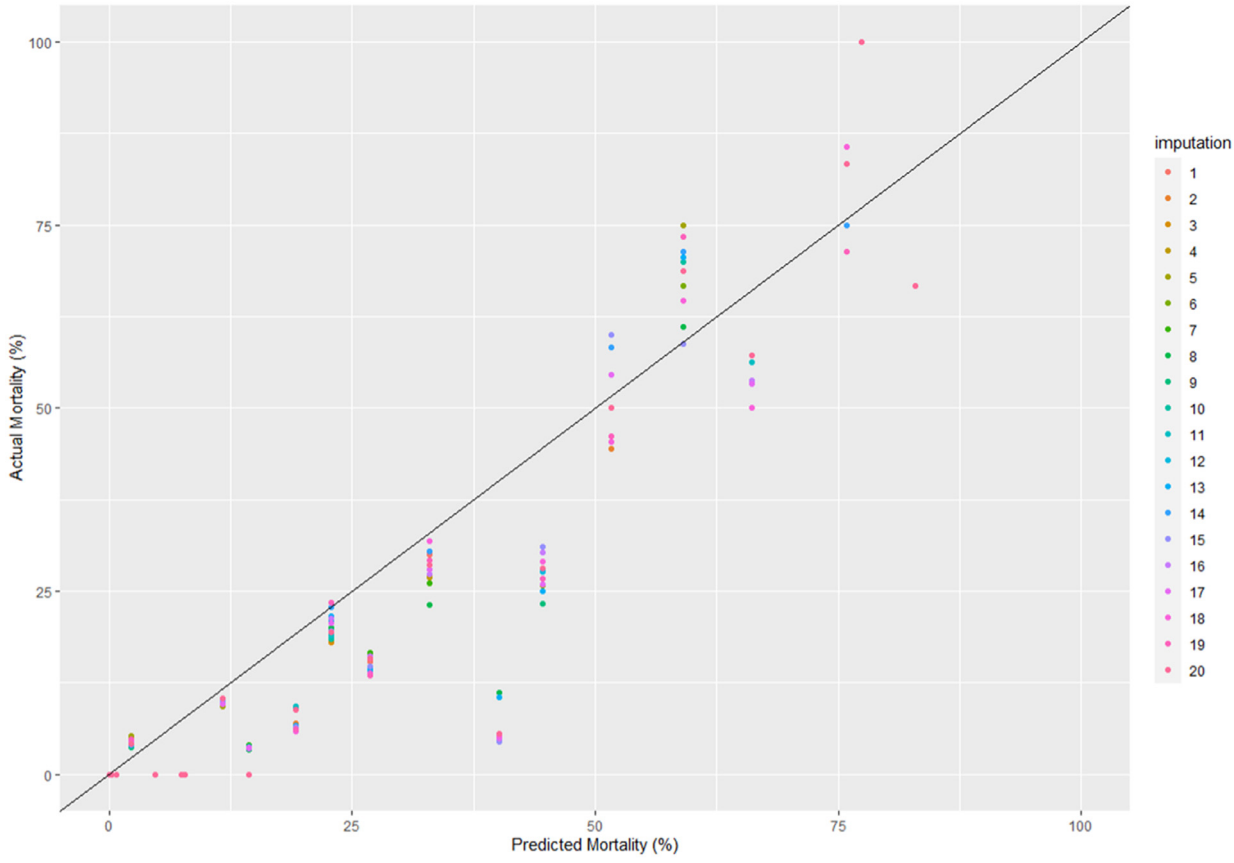

FIGURE 1. Predicted (4C Mortality Score) vs. actual mortality for all imputed datasets.

socioeconomic status, and COVID-19 outcomes highlights the necessity for including under-represented minorities in the validation of risk models.[22-24]

There are notable limitations to this study, including the limited generalizability of a validation performed in a single, tertiary care center. The absence of dementia and "chronic neurological disease" may have yielded lower risk scores for patients with 0-1 other comorbid conditions. Additionally, the original 4C Mortality Score used clinician-defined obesity, whereas this study used a BMI of > 30 kg/m^2. We anticipate that the effect of missing comorbidities is opposed by the increased sensitivity of manual chart review in identifying comorbidities, mitigating any significant effect on model calibration, especially considering that comorbidities account for a small portion of the 4C Mortality Score. The primary outcome or the original 4C Mortality Score was 30-day in-hospital mortality, as compared to mortality during admission or within 30 days of discharge. As the intent of our validation of the 4C Mortality Score was to evaluate its clinical utility, we chose to include patients who died within 30 days of discharge under the assumption that they were likely to have died of complicated related to their index admission; that the average time from index admission to death was 13.9 days appears to support this assumption. Given these limitations, we believe the overall conclusion is still appropriate and reinforces the importance of clinicians and public health leaders thoughtfully applying the 4C Score to current specific scenarios.

The 4C Mortality Score is a clinically useful tool to assess risk for COVID-19 mortality as validated in an early 2020 United States urban tertiary care center. The validation of the 4C Mortality Score in a majority Black population is important for providers in many United States urban and rural centers to confidently use this tool for clinical decision making. The score is particularly useful for emergency department triage, as well as patient-specific treatment decision making (e.g., monoclonals, oral antivirals, remdesivir dosing, etc.) for scenarios in which treatment regimens are not clearly defined in national guidelines. Given the vaccination efforts and improved treatment options since this study cohort, the 4C Mortality Score should be applied thoughtfully by clinicians, carefully accounting for differences between an individual patient's clinical scenario and the populations and timing of our validation cohort. Further external validation of COVID-19 risk prediction tools in cohorts including vaccinated and unvaccinated patients receiving more recent treatment regimens would be of great value to clinicians. Ultimately, the 4C Mortality Score provides one of the best evidence-based

approaches to date for patient triage, treatment choice and duration, and discharge planning to best serve patients afflicted by COVID-19.

FUNDING

None.

CONFLICTS OF INTEREST

The authors have no conflicts of interest to disclose.

SUPPLEMENTARY MATERIALS

Supplementary material associated with this article can be found in the online version at https://doi.org/10.1016/j.amjms.2022.04.030.

REFERENCES

1. **Liang W, Liang H, Ou L, et al.** Development and validation of a clinical risk score to predict the occurrence of critical illness in hospitalized patients with COVID-19. *JAMA Intern Med*. 2020;180(8):1081–1089.
2. **Xie J, Hungerford D, Chen H, et al.** Development and external validation of a prognostic multivariable model on admission for hospitalized patients with COVID-19. medRxiv. 2020. doi:10.1101/2020.03.28.20045997.
3. **Woo SH, Rios-Diaz AJ, Baram M, et al.** Development and validation of a web-based severe COVID-19 risk prediction model. *Am J Med Sci*. 2021;362(4):355–362.
4. **Knight SR, Ho A, Pius R, et al.** Risk stratification of patients admitted to hospital with covid-19 using the ISARIC WHO clinical characterisation protocol: development and validation of the 4C mortality score. *BMJ*. 2020;370:m3339.
5. **Gupta RK, Harrison EM, Ho A, et al.** Development and validation of the ISARIC 4C Deterioration model for adults hospitalised with COVID-19: a prospective cohort study. *Lancet Respir Med*. 2021;9(4):349–359.
6. **van Dam PMEL, Zelis N, van Kuijk SMJ, et al.** Performance of prediction models for short-term outcome in COVID-19 patients in the emergency department: a retrospective study. *Ann Med*. 2021;53(1):402–409.
7. **Neto FL, Marino LO, Torres A, et al.** Community-acquired pneumonia severity assessment tools in patients hospitalized with COVID-19: a validation and clinical applicability study. *Clin Microbiol Infect*. 2021. S1198-743X(21)00136-1.
8. **Covino M, De Matteis G, Burzo ML, et al.** Predicting in-hospital mortality in COVID-19 older patients with specifically developed scores. *J Am Geriatr Soc*. 2021;69(1):37–43.
9. Rubin DB. Multiple Imputation for Nonresponse in Surveys. New York: John Wiley and Sons; 2004.
10. **Lassau N, Ammari S, Chouzenoux E, et al.** Integrating deep learning CT-scan model, biological and clinical variables to predict severity of COVID-19 patients. *Nat Commun*. 2021;12:634.
11. **Verma AA, Hora T, Jung HY, et al.** Characteristics and outcomes of hospital admissions for COVID-19 and influenza in the Toronto area. *CMAJ*. 2021;193(12):E410–E418.
12. **Tenforde MW, Self WH, Adams K, et al.** Association Between mRNA Vaccination and COVID-19 Hospitalization and Disease Severity. *JAMA*. 2021;326(20):2043–2054.
13. **Gupta A, Gonzalez-Rojas Y, Juarez E, et al.** Early treatment for Covid-19 with SARS-CoV-2 neutralizing antibody sotrovimab. *N Engl J Med*. 2021;385:1941–1950.
14. **Gottlieb RL, Vaca CE, Paredes R, et al.** Early remdesivir to prevent progression to severe Covid-19 in outpatients. *N Engl J Med*. 2022;386:305–315.
15. **Beigel JH, Tomashek KM, Dodd LE, et al.** Remdesivir for the treatment of Covid-19 — final report. *N Engl J Med*. 2020;383:1813–1826.
16. **Group The RECOVERY Collaborative.** Dexamethasone in hospitalized patients with covid-19. *N Engl J Med*. 2021;384:693–704.
17. **Group The RECOVERY Collaborative.** Tocilizumab in patients admitted to hospital with COVID-19 (RECOVERY): a randomised, controlled, open-label, platform trial. *Lancet*. 2021;397:1637–1645.
18. **Marconi VC, Ramanan AV, de Bono S, et al.** Efficacy and safety of baricitinib for the treatment of hospitalised adults with COVID-19 (COV-BARRIER): a randomised, double-blind, parallel-group, placebo- controlled phase 3 trial. *Lancet Resp*. 2021;9:1407–1418.
19. **Perkins GD, Ji C, Connolly BA, Couper K, et al.** Effect of noninvasive respiratory strategies on intubation or mortality among patients with acute hypoxemic respiratory failure and COVID-19: the RECOVERY-RS randomized clinical trial. *JAMA*. 2022;327(6):546–558.
20. **Ospina-Tascón GA, Calderón-Tapia LE, García AF, et al.** Effect of high-flow oxygen therapy vs conventional oxygen therapy on invasive mechanical ventilation and clinical recovery in patients with severe COVID-19: a randomized clinical trial. *JAMA*. 2021;326(21):2161–2171.
21. **Schaefer JW, Riley JR, Li M, et al.** Comparing reliability of ICD-10-based COVID-19 comorbidity data to manual chart review, a retrospective cross-sectional study. *J Med Virol*. 2022;94(4):1550–1557.
22. **Magesh S, John D, Li WT, et al.** Disparities in COVID-19 outcomes by race, ethnicity, and socioeconomic status: a systematic-review and meta-analysis. *JAMA Netw. Open.*. 2021;4:(11) e2134147.
23. **Li AY, Hannah TC, Chouhdri TF, et al.** Multivariate analysis of black race and environmental temperature on COVID-19 in the US. *Am J Med Sci*. 2020;360(4):348–356.
24. **Cheney-Peters DR, Lee CY, Mitsuhashi S, et al.** Association of race/ethnicity and socioeconomic status with COVID-19 30-day mortality at a Philadelphia medical center using a retrospective cohort study. *J Med Virol*. 2022;94(3):906–917.

Submitted August 22, 2021; accepted April 22, 2022.

Corresponding author at: Alan A. Kubey, MD, Division of Hospital Medicine, Thomas Jefferson University Hospital, 833 Chestnut Street, Suite 701, Philadelphia, PA 19107, USA (E-mail: Alan.kubey@jefferson.edu).