Title

External validation of the 4C Mortality Score for hospitalised patients with COVID-19 in the RECOVER network

Journal

BMJ Open, 12(4)

ISSN

2044-6055

Authors

Gordon, Alexandra June
Govindarajan, Prasanthi
Bennett, Christopher L
et al.

Publication Date

2022-04-01

DOI

10.1136/bmjopen-2021-054700

Peer reviewed

BMJ Open

External validation of the 4C Mortality Score for hospitalised patients with COVID-19 in the RECOVER network

Alexandra June Gordon ,[1] Prasanthi Govindarajan,[1] Christopher L Bennett,[1,2] Loretta Matheson,[1] Michael A Kohn,[1,3] Carlos Camargo,[4] Jeffrey Kline [5]

To cite: Gordon AJ, Govindarajan P, Bennett CL, *et al*. External validation of the 4C Mortality Score for hospitalised patients with COVID-19 in the RECOVER network. *BMJ Open* 2022;**12**:e054700. doi:10.1136/bmjopen-2021-054700

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2021-054700).

Received 10 July 2021
Accepted 28 March 2022

[1]Emergency Medicine, Stanford University School of Medicine, Stanford, California, USA
[2]Epidemiology, Stanford University School of Medicine, Stanford, California, USA
[3]Epidemiology and Biostatistics, University of California San Francisco, San Francisco, California, USA
[4]Emergency Medicine, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts, USA
[5]Emergency Medicine, Wayne State University School of Medicine, Detroit, Michigan, USA

Correspondence to
Dr Jeffrey Kline;
jkline@wayne.edu

ABSTRACT

Objectives Estimating mortality risk in hospitalised SARS-CoV-2+ patients may help with choosing level of care and discussions with patients. The Coronavirus Clinical Characterisation Consortium Mortality Score (4C Score) is a promising COVID-19 mortality risk model. We examined the association of risk factors with 30-day mortality in hospitalised, full-code SARS-CoV-2+ patients and investigated the discrimination and calibration of the 4C Score. This was a retrospective cohort study of SARS-CoV-2+ hospitalised patients within the RECOVER (REgistry of suspected COVID-19 in EmeRgency care) network.

Setting 99 emergency departments (EDs) across the USA.

Participants Patients ≥18 years old, positive for SARS-CoV-2 in the ED, and hospitalised.

Primary outcome Death within 30 days of the index visit. We performed logistic regression analysis, reporting multivariable risk ratios (MVRRs) and calculated the area under the ROC curve (AUROC) and mean prediction error for the original 4C Score and after dropping the C reactive protein (CRP) component.

Results Of 6802 hospitalised patients with COVID-19, 1149 (16.9%) died within 30 days. The 30-day mortality was increased with age 80+ years (MVRR=5.79, 95% CI 4.23 to 7.34); male sex (MVRR=1.17, 1.05 to 1.28); and nursing home/assisted living facility residence (MVRR=1.29, 1.1 to 1.48). The 4C Score had comparable discrimination in the RECOVER dataset compared with the original 4C validation dataset (AUROC: RECOVER 0.786 (95% CI 0.773 to 0.799), 4C validation 0.763 (95% CI 0.757 to 0.769). Score-specific mortalities in our sample were lower than in the 4C validation sample (mean prediction error 6.0%). Dropping the CRP component from the 4C Score did not substantially affect discrimination and 4C risk estimates were now close (mean prediction error 0.7%).

Conclusions We independently validated 4C Score as predicting risk of 30-day mortality in hospitalised SARS-CoV-2+ patients. We recommend dropping the CRP component of the score and using our recalibrated mortality risk estimates.

Strengths and limitations in this study

- In this first study using a national US sample of patients who tested positive for SARS-CoV-2 and were hospitalised through emergency departments, our results confirmed the previous findings that older age, comorbidities, body mass index≥40 kg/m^2, higher respiratory rate and lower oxygen saturation were associated with 30-day mortality.
- We also observed that the arrival to the emergency care setting from a nursing home was associated with increased mortality.
- We independently validated 4C Mortality Score as predicting risk of 30-day mortality in hospitalised SARS-CoV-2+ patients.
- We recommend dropping the C reactive protein component of the score and using our recalibrated mortality risk estimates when estimating the 30-day mortality in hospitalised patients who test positive for SARS-CoV-2.

INTRODUCTION

The COVID-19 pandemic has placed tremendous strain on emergency and critical care resources in hospitals worldwide.[1–3] To prepare the healthcare systems for the surges, several studies have developed prediction models to assess mortality risk in patients hospitalised with COVID-19. These studies identified the following risk factors for mortality or critical care admission: age, sex, comorbid conditions, laboratory values and vital signs.[4–16]

In a systematic review that evaluated many of these risk prediction studies using the prediction model risk of bias assessment tool (PROBAST), Wynants *et al* concluded that many of the current risk models may be misleading.[10] However, the authors' analysis suggested that one COVID-19 mortality prediction model, the Coronavirus Clinical Characterisation Consortium (4C) Mortality Score, which was built on a large UK dataset, had relatively low risk of bias in most domains by the PROBAST criteria. The 4C Mortality Score includes eight variables: age, sex, respiratory rate, oxygen saturation, number of comorbidities, level of consciousness, blood urea nitrogen and C reactive protein (CRP) (see table 1).[17] While there has been

Table 1 Point assignment for 4C Mortality Score

Age Groupgroup (years)	
18–49	
50–59	+2
60–69	+4
70–79	+6
80+	+7
Sex at birth	
Female	
Male	+1
Comorbidities*	
0	
1	+1
2 +	+2
Respiratory rate (breaths/min)	
<20	
20–29	+1
≥30	+2
Oxygen saturation, room air (%)	
≥92	
<92	+2
Altered mental status†	
No	
Yes	+2
Blood urea nitrogen (mg/dL)‡	
<20	
20–39	+1
≥40	+3
C reactive protein (mg/dL)	
<5.0	
5.0–9.9	+1
≥10.0	+2

*Comorbidities: High body mass index, cancer, chronic cardiac disease, chronic pulmonary disease, diabetes, liver disease, kidney disease.
†Altered mental status is patient-reported symptom, whereas 4C used Glasgow Coma Scale <15.
‡Blood urea nitrogen and C reactive protein units converted from 4C.
4C Mortality Score, Coronavirus Clinical Characterisation Consortium Mortality Score.

continued interest in the development of prediction models for COVID-19, the 4C Mortality Score represented one of the first with a low risk of bias and therefore a good candidate for verification in other populations.

In this study, we investigated the risk of 30-day mortality in hospitalised SARS-CoV-2+ patients within the RECOVER (REgistry of suspected COVID-19 in EmeRgency care) network.[18] In a large cohort of SARS-CoV-2+ patients hospitalised from 99 US emergency departments (EDs), we determined the relation of demographic and clinical factors with 30-day mortality and investigated the discrimination and calibration of the 4C Mortality Score with and without the CRP value.

METHODS

In this retrospective cohort study, we included patient-level data from the RECOVER Network, a national registry of patients who were tested for SARS-CoV-2 during their ED visit. We restricted the analysis to full code status patients ≥18 years old who tested positive for SARS-CoV-2 and were hospitalised from the ED.[18] The study was approved or deemed exempted by the Institutional Review Boards of all participating sites.

Data source

We obtained data from 40 medical centres representing 99 EDs from 27 US states and the District of Columbia. Data were collected using a REDCap data collection form that was distributed to the ED sites during the study period (March 2020–September 2020); our data were downloaded from the registry in December 2020. The REDCap form (online supplemental appendix A) had seven sections and 204 questions, which generated 360 data elements. Variables reflected a combination of routinely collected information (eg, patient demographics, medical history, vital signs and diagnostic test results), patient-reported symptoms and risk exposures, clinical outcomes (eg, admission, therapies, death) and those deemed important by the RECOVER Network steering committee. After creation, but prior to launch, the data form was piloted at 19 sites and refined. For additional section details and the questions, please refer to the data collection form in the online supplemental material. The data were obtained from the electronic healthcare record using a combination of electronic download for routinely collected, coded variables (eg, age, vital signs and laboratory values), supplemented by chart review by research personnel, using methods previously described.[18]

Patient and public involvement

Patients were not involved in the development of our work, setting the research question or determining the outcome measures. This applies to both the RECOVER network and the work presented here. Given the nature and limitations of emergency care during the COVID-19 pandemic, it was not appropriate or possible to involve patients or the public in the design, or conduct, or reporting or dissemination plans of our work.

Study variables

We analysed patient characteristics such as demographics, vital signs, symptoms, risks for infection, comorbidities and medications. Following the 4C Mortality Score, we categorised the patients into five age groups (18–49, 50–59, 60–69, 70–79, 80+). The US standard ethnicity

(Hispanic/Latinx yes/no) and race categories were combined into eight categories (Hispanic/Latinx (H/L), non-H/L white, non-H/L African American/black, non–H/L Asian, non-H/L Native American/Alaskan Native, non-H/L Pacific Islander/Native Hawaiian, non-H/L mixed and Unknown). In the analysis, we combined non–H/L Native American (0.2%), non-H/L Pacific Islander (0.2%), non-H/L mixed, other and unknown into a single group (12.8%).

All included patients had a positive reverse transcriptase PCR test (RT-PCR) test for SARS-CoV-2. Almost all the tests were performed during the ED visit, but we also included patients who had a test in a physician's office or urgent care immediately prior to the ED visit. We excluded patients whose 30-day vital status could not be ascertained, those who died in the ED before vital signs were recorded and those who did not have full code status.

Study outcome

The primary outcome was death within 30 days of the index visit. The 4C Mortality Score's predictive accuracy was measured by the area under the ROC curve (AUROC) and mean prediction error.

Data preparation and statistical analysis

For comparison with the other cohorts, we report the median and IQR of continuous variables—both in the entire cohort and in the subgroup who died within 30 days—and compared the median values using the rank sum test. We performed univariable analysis on 26 independent variables that were included on the data collection form using complete (non-missing) data and reporting risk ratios for 30-day mortality. For risk ratio reporting of continuous variables, we chose category boundaries based on the 4C Mortality Score (age, respiratory rate, oxygen saturation, blood urea nitrogen (BUN) and CRP) or other published mortality prediction models (body mass index (BMI), creatinine, total bilirubin).[19]

We selected variables for our multivariable logistic regression model based on the 4C Mortality Score, other prior studies and clinical judgement. The RECOVER dataset has complete data (<1.5% missing) on most variables, with the exception of CRP, BMI, bilirubin and smoking status. For our multivariable analysis, we used imputed values for missing data using Stata's implementation of the data augmentation algorithm.[20] We report multivariable risk ratios with 95% CIs. Statistical analysis

Table 2 Patient characteristics, clinical characteristics and 30-day mortality of SARS-CoV-2+ patients hospitalised from the emergency department

	Total	Median (IQR)	Deaths	Median (IQR)
Overall	6802		1149	
Key clinical measures				
Age, in years	6802	64 (52–75)	1149	74 (64–84)
Oxygen saturation, room air (%)	6802	92 (87–95)	1149	86 (76–93)
C reactive protein (mg/dL)	4163	10.5 (4.8–18.9)	643	17.3 (9.4–25.7)
BMI	6058	28.7 (24.7–34)	970	28 (24.1–33.2)
Other vital signs				
Temperature (°C)	6801	37.2 (36.7–37.9)	1149	37.1 (36.6–37.8)
Heart rate (beats/min)*	6800	97 (84–110)	1148	98 (83–112)
Systolic blood pressure (mm Hg)	6802	130 (116–146)	1149	127 (112–146)
Respiratory rate (breaths/min)	6797	20 (18–23)	1148	20 (18–26)
Other blood tests				
White cell count (10^9/L)	6767	7.1 (5.3–9.9)	1140	8.9 (6.2–12)
Haemoglobin (g/L)	6769	130 (120–146)	1142	130 (110–145)
Platelets (10^9 cells/L)	6760	209 (160–275)	1140	201.5 (149–270)
Sodium (mEq/L)	6338	136 (133–139)	1039	137 (133–142)
Potassium (mEq/L)	6743	4.1 (3.8–4.6)	1141	4.4 (4–5.1)
Blood urea nitrogen (mg/dL)	6706	18 (12–31)	1131	32 (19–54)
Creatinine (mg/dL)	2832	1 (0.8–1.4)	214	1.3 (1–2.1)
Total bilirubin (mg/dL)	6124	0.6 (0.4–0.8)	1051	0.6 (0.4–0.9)

P value for rank sum test comparison of died versus survived are p<=0.003 except heart rate.
*P value=0.62.
BMI, body mass index.

Table 3 Effect of patient and clinical characteristics on 30-day mortality of SARS-CoV-2+ patients hospitalised from the emergency department

	Total	% of sample	Deaths	30-day mortality (% died)	Relative risk	Multivariable relative risk (95% CI)
Overall	6802		1149	16.9		
In both 4C Score and multivariable model						
Age group (years)						
18–49	1413	20.8	63	4.5	Reference	Reference
50–59	1272	18.7	108	8.5	1.90	1.66 (1.18 to 2.14)
60–69	1690	24.9	263	15.6	3.49	2.84 (2.09 to 3.58)
70–79	1302	19.1	293	22.5	5.05	4.03 (2.98 to 5.08)
80+	1125	16.5	422	37.5	8.41	5.79 (4.23 to 7.34)
Sex at birth						
Female	2980	43.8	466	15.6	Reference	Reference
Male	3822	56.2	683	17.9	1.14	1.17 (1.05 to 1.28)
Respiratory rate (breaths/min)						
<20	2896	42.6	384	13.3	Reference	Reference
20–29	3282	48.3	567	17.3	1.30	1.12 (1 to 1.24)
≥30	619	9.1	197	31.8	2.40	1.66 (1.42 to 1.9)
Oxygen saturation, room air (%)						
≥92	4017	59.1	364	9.1	Reference	Reference
<92	2785	40.9	785	28.2	3.11	2.32 (2.06 to 2.58)
C reactive protein (mg/dL)						
<5.0	1064	25.6	64	6.0	Reference	Reference
5.0–9.9	947	22.8	108	11.4	2.23	1.23 (1.01 to 1.45)
≥10.0	2152	51.7	471	21.9	4.52	1.7 (1.44 to 1.95)
In 4C Score (only towards comorbidity count) and multivariable model						
*BMI						
18.5–<40	5227	86.3	823	15.7	Reference	Reference
<18.5	175	2.9	41	23.4	1.49	0.96 (0.72 to 1.2)
≥40	656	10.8	106	16.2	1.03	1.44 (1.23 to 1.64)
*Cancer	547	8.0	88	16.1	0.95	0.81 (0.67 to 0.96)
*Chronic cardiac disease (any of below)	1170	17.2	277	23.7	1.53	1.06 (0.93 to 1.19)
Atrial fibrillation	542	8.0	149	27.5	1.72	
Heart disease	382	5.6	69	18.1	1.07	
Heart failure	608	8.9	142	23.4	1.44	
*Chronic pulmonary disease (any of below)	529	7.8	112	21.2	1.28	1.06 (0.89 to 1.24)
COPD	433	6.4	91	21.0	1.27	
Bronchiectasis	17	0.3	3	17.6	1.04	
Other lung disease	128	1.9	27	21.1	1.25	
Pulmonary fibrosis	18	0.3	5	27.8	1.65	
*Diabetes	2079	30.6	357	17.2	1.02	0.97 (0.87 to 1.07)

Continued

Table 3 Continued

	Total	% of sample	Deaths	30-day mortality (% died)	Relative risk	Multivariable relative risk (95% CI)
*Liver disease (total bilirubin≥2.0)	175	2.9	52	29.7	1.77	1.56 (1.24 to 1.88)
*Kidney disease (creatinine ≥1.2 or BUN ≥40)	1877	28.1	515	27.4	2.14	1.58 (1.42 to 1.74)
In 4C Score only						
Comorbidities, among seven with*						
0	2632	38.7	300	11.4	Reference	
1	2207	32.5	418	18.9	1.08	
2 +	1963	28.9	431	22.0	1.12	
Altered mental status	957	14.1	162	16.9	1.00	
Blood urea nitrogen (mg/dL)						
<20	3715	55.4	297	8.0	Reference	
20–39	1771	26.4	390	22.0	2.75	
≥40	1220	18.2	444	36.4	4.56	
In multivariable model only						
Race/ethnicity						
White, non-H/L	1652	24.3	323	19.6	Reference	Reference
Asian, non-H/L	234	3.4	41	17.5	0.90	1.05 (0.77 to 1.33)
Black, non-H/L	2286	33.6	362	15.8	0.81	1 (0.87 to 1.13)
Hispanic/Latinx (H/L)	1759	25.9	228	13.0	0.66	0.96 (0.82 to 1.1)
Other or unknown	871	12.8	195	22.4	1.14	1.3 (1.11 to 1.49)
Resides in nursing home or assisted living	703	10.3	256	36.4	2.49	1.29 (1.1 to 1.48)
Smoker	447	7.5	48	10.7	0.70	1.02 (0.82 to 1.23)
Asthma	581	8.5	68	11.7	0.67	0.89 (0.71 to 1.06)

Where missing is over 1.5%: C reactive protein=38.8%; BMI=10.9%; Total bilirubin=10.0%; Smoker=12.7%.
'Relative risk' is the risk of death relative to the reference if indicated, otherwise to not having the risk factor.
BMI, body mass index; COPD, chronic obstructive pulmonary disease.

was performed using SAS Enterprise Guide V.8.3 and Stata/SE V.16.1.[21]

We replicated the 4C Mortality Score described by Knight *et al* with one modification.[17] Since we did not have a variable for Glasgow Coma Score or confusion on examination, we used the symptom 'altered mental status or confusion' instead. In addition to the full score, we tested a modified score dropping CRP, which was missing in 39% of the records. We evaluated discrimination and calibration using nine score categories available from Knight *et al* (0–2, 3–4, 5–6, 7–8, 9–10, 11–12, 13–14, 15–16, ≥17). We used the mortality reported in the 4C validation dataset as our predicted risks for comparison with observed mortality. For reporting, we pooled the results into the four risk groups defined by Knight *et al*: Low: 0–3; Intermediate: 4–8; High: 9–14; and Very High ≥15. The AUROC was calculated with 95% CI using the DeLong method.[22] Calibration was assessed using a standard calibration table, mean prediction error and the square root of both the calibration error and the Brier Score. We also used a modified Bland-Altman-style calibration plot.[23]

RESULTS

Of 26 914 patients in the first version of the registry, 6822 met the inclusion criteria for this analysis (≥18 years old, SARS-CoV-2+, hospitalised from the ED, full code status). We excluded 11 who were missing vital status at 30 days

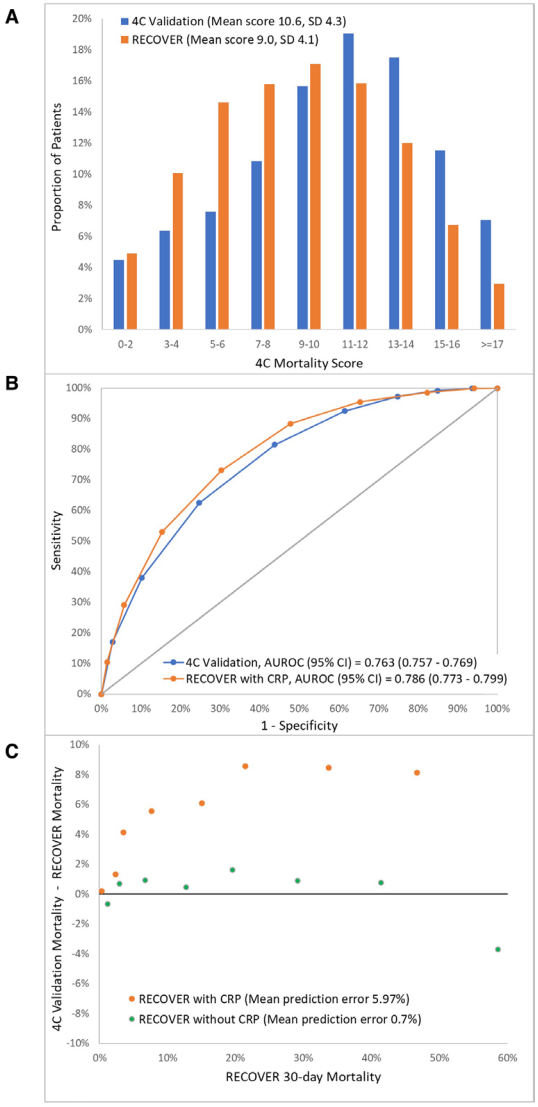

Figure 1 Comparison of 4C (Coronavirus Clinical Characterisation Consortium) validation and RECOVER datasets. (A) 4C Mortality Scores were lower in the RECOVER dataset than in the original 4C validation dataset. (B) ROC curves for the 4C Mortality Score (categorised into the nine ranges from (A)) in the 4C validation dataset and the RECOVER dataset. (C) Calibration plot (modified Bland-Altman) showing prediction error versus observed mortality for the 4C Mortality Score with and without the C reactive protein (CRP) component. Points from left to right are in the 4C Mortality Score ranges shown in figure (A) from left to right. AUROC, area under the ROC curve; RECOVER, REgistry of suspected COVID-19 in EmeRgency care.

and 9 who died in the ED prior to vital signs, leaving 6802 in the analysis cohort. Of the 6802, 1149 (16.9%) died within 30 days. The median age of patients in the cohort was 64 years (IQR 52–75); 56.2% were male; and 61.4% had at least one comorbid condition. Of note, the median oxygen saturation was 92% (IQR 87%–95%) overall and 86% (IQR 76%–93%) in those who died (p<0.0001) (table 2).

Of the demographic risk factors, age group, male sex and residence in a nursing home/assisted living facility were the principal mortality predictors (table 3). In the multivariable model, age 80+ years increased 30-day mortality risk by a factor of 5.79 (95% CI 4.23 to 7.34); male sex increased it by 1.17 (95% CI 1.05 to 1.28); and nursing home/assisted living facility residence increased it by 1.29 (95% CI 1.1 to 1.48). On a univariable basis, Hispanic ethnicity and smoking status were associated with lower mortality risk, but after adjusting for other variables, including the younger age of Hispanics and smokers, the risk ratio for mortality for Hispanic ethnicity was 0.96 (95% CI 0.82 to 1.1) and for smoking was 1.02 (95% CI 0.82 to 1.23).

In the univariable analysis, extreme obesity (BMI≥40) did not increase risk, but after adjustment for age, sex and other comorbidities, the risk ratio for BMI≥40 was 1.44 (95% CI 1.23 to 1.64). In addition to obesity, the multivariable analysis (table 3) showed that other comorbidities associated with increased risk of death were liver disease as indicated by a total bilirubin ≥2.0 mg/dL and kidney disease as indicated by a creatinine ≥1.2 mg/dL or BUN ≥40. Asthma and diabetes were not significant risk factors. Patients who arrived from a nursing home had an increased risk of mortality (risk ratio 1.29, 95% CI 1.1 to 1.48).

Table 3 also shows that increase in respiratory rate, decrease in oxygen saturation and increase in CRP each corresponded with an increase in mortality.

Compared with the 4C validation dataset from Knight *et al*, the mean 4C Mortality Scores were lower in our dataset (mean score 9.0 vs 10.6) (figure 1A). The AUROC from the RECOVER dataset was comparable to that of the original 4C validation dataset. Using nine 4C score categories, the AUROC from the RECOVER dataset was not substantially different than the AUROC from the original 4C validation dataset (AUROC: RECOVER 0.786 (95% CI 0.773 to 0.799) versus 4C validation 0.763 (95% CI 0.757 to 0.769) (figure 1B). Our observed category-specific mortalities were lower than those in the 4C validation dataset. Using the mortalities from the 4C validation dataset would have overestimated risk by 6.0% on average (Mean prediction error 6.0%; √Calibration error 0.066; and √Brier Score 0.350) (table 4, figure 1C).

Dropping CRP from the 4C Mortality Score reduced the scores overall (mean score 7.7) but did not substantially change discrimination (AUROC 0.776, 95% CI 0.762 to 0.790). Dropping the CRP component did affect calibration. The category-specific mortalities in our dataset were now close to those in the 4C validation dataset. Using the mortalities from the 4C validation dataset would have misestimated risk by 0.7% on average (Mean prediction error 0.7%, √Calibration error 0.017 and √Brier Score 0.346).

DISCUSSION

In this analysis of multicentre data from the RECOVER network, our results confirmed several previous findings for risk factors for COVID-19 mortality, including older age, comorbidities, BMI≥40 kg/m^2, higher respiratory rate and lower oxygen saturation.[4–9 11–14] In addition, as reported by Graselli *et al* in critically ill patients,

Table 4 Comparison of observed mortality by 4C mortality risk group for recover dataset of SARS-CoV-2+ patients hospitalised from the emergency department

	RECOVER dataset with CRP			RECOVER dataset without CRP		
4C mortality risk group (score range)	Mortality predicted by 4C*	Observed mortality	Prediction error	Mortality predicted by 4C*	Observed mortality	Prediction error
Overall	22.9%	16.9%	6.0%	17.6%	16.9%	0.7%
Low (0–3 points)	1.4%	0.8%	0.6%	1.2%	1.6%	–0.5%
Intermediate (4–8 points)	9.7%	5.3%	4.4%	9.5%	8.7%	0.8%
High (9–14 points)	29.9%	22.3%	7.6%	28.6%	27.4%	1.2%
Very high (≥15 points)	60.2%	50.6%	9.6%	56.8%	58.2%	–1.4%
	AUROC		√Calibration error		√Brier Score	
4C validation	0.763 (0.757–0.769)					
RECOVER dataset with CRP	0.786 (0.773–0.799)		0.066		0.350	
RECOVER dataset without CRP	0.776 (0.762–0.79)		0.017		0.346	

*The mortality predicted by 4C is constant for each individual score, but when the scores are grouped into ranges (as they are here), the predicted mortality varies based on the proportion of patients from the test dataset with each individual score within the range.
CRP, C reactive protein; RECOVER, REgistry of suspected COVID-19 in EmeRgency care.

we observed that male sex is predictive of mortality.[7] We also observed the expected, but previously unquantified finding that arrival to the emergency care setting from a nursing home was associated with increased mortality. While this has not been specifically mentioned in other studies, Ferrando-Vivas *et al* found that functional dependence was related to mortality (HR 1.425).[5]

In the RECOVER network, COVID-19 positivity was higher among Hispanic patients when compared with non-Hispanics, but the adjusted mortality among hospitalised Hispanic patients is similar to hospitalised non-Hispanic whites.[24] Similarly, Mackey *et al* reported that hospitalisations for COVID-19 among those who identify their ethnicity as Hispanic were proportionately higher than for their non-Hispanic white counterparts but the case fatality rate was similar between Hispanic and non-Hispanic patients.[25]

We also found that the comorbid conditions such as liver disease defined as elevated total bilirubin ≥2.0 and kidney disease defined as creatinine ≥1.2 mg/dL or BUN ≥40 had an independent association with 30-day mortality in hospitalised SARS-CoV-2+ patients. Surprisingly, previous studies and our results did not establish diabetes as a significant risk factor.[26–28] Our findings on the association of smoking with 30-day mortality did not concur with previous studies. Smoking and cumulative smoking exposure were predictive of mortality in previous studies,[26] but we did not find a statistically significant association after controlling for other variables. Finally, among the clinical variables, tachypnoea (respiratory rate ≥20) and hypoxaemia (oxygen saturation <92%) were significant predictors of mortality. Zhao *et al* reported higher odds of mortality (adjusted OR 4.8) for an oxygen saturation <92%.[13]

Given the multiplicity of variables associated with 30-day mortality, clinicians need a simple score to better predict short-term mortality. The 4C Mortality Score is one such score and it performed well in our dataset. Discrimination was excellent, and calibration was also good, although using the category-specific mortalities from the 4C validation dataset would have overestimated risk. CRP was missing in 39% of the records in our study, so we examined the performance of the 4C Mortality Score without the CRP component. Discrimination remained good, and the category-specific risks from the 4C validation were accurate. When CRP was removed from the score, many patients with high CRP values moved into a lower risk category. Those patients who remained with high 4C Mortality Scores despite removal of CRP died at a higher rate than those whose risk score decreased, but those with high CRP values who moved to a lower risk group had higher mortality than the average for their new lower risk group. This might be referred to as stage migration effect. When the high CRP patients moved from the *very-high-risk* group to the *high-risk* group, the average mortality went up in both groups. Based on our observations, we suggest using the 4C Mortality Score without the CRP component, but recalibrating risk estimates as per our table 4 or online supplemental table A. Using category-specific risks as opposed to the four risk groups (low, intermediate, high, very high) is preferred because it does not assume the distribution across the risk groups is the same in different populations. This modified 4C Mortality Score could assist with triage decisions, to inform patients

and their family members of prognostic information, and to help with forecasting of resource utilisation in the hospital.

The nature of the COVID-19 pandemic greatly accelerated the timeline of related research and has resulted in rapid changes to practice patterns and patient presentation. At the time of this study, the 4C Mortality Score was among the most promising risk evaluation tools and had been identified as having a low likelihood of bias. Since the inception of our study to validate this score, many other systems have been proposed. These have been developed in a variety of different patient populations using a wide range of methods.[27–35] Some models have been independently assessed and performance varies.[36] Updates to a systematic review of prediction models continue to identify the prognostic 4C Mortality Score as among the most promising,[37] suggesting that attempts to validate and calibrate this and other existing risk estimation models could aid providers in the evaluation of the many available scoring systems for patients with COVID-19 disease.

Limitations

This is a national study of hospitalised SARS-CoV-2+ patients. The large sample size, the number and diversity of the participating sites, and a comprehensive list of data elements are major strengths of this study. However, some sites contributed more SARS-CoV-2+ patients than others. We did see regional differences in 30-day mortality, but these did not affect the risk ratios. As noted above, CRP was missing in almost 39% of patients.

Additional limitations are related to the nature of the COVID-19 pandemic and the changes in patient population and clinical practices that have occurred over time. The data in this study represent a time period early in the pandemic (on or before September 2020) and thus may not fully account for practice changes. However, these data align with the time period of the RECOVERY trial, which introduced the main practice change affecting mortality (use of dexamethasone) in February 2021.[38]

CONCLUSIONS

We conclude that among SARS-CoV-2+ hospitalised patients, older patients with comorbid conditions and those with hypoxaemia at the time of presentation have a very high risk of dying within 30 days. We independently validate the 4C Mortality Score as predicting risk of death in hospitalised SARS-CoV-2+ patients, but we recommend dropping the CRP component of the score and using our recalibrated mortality risk estimates.

Twitter Alexandra June Gordon @ajunegordon

Collaborators The RECOVER Investigators Group: Thomas Aufderheide, MD, Medical College of Wisconsin, Milwaukee, WI. Joshua Baugh, MD, Massachusetts General Hospital, Boston, MA, USA. David Beiser, MD, University of Chicago School of Medicine, Chicago, IL, USA. Joseph Bledsoe, MD, Healthcare Delivery Institute, Intermountain Healthcare, Salt Lake City, UT, USA. Edward Castillo, PhD, University of California at San Diego, San Diego, CA, USA. Makini Chisholm-Staker, MD, Mt. Sinai School of Medicine, New York, NY, USA. D. Mark, Courtney, MD, University of Texas Southwestern, Dallas, TX, USA. Elizabeth Goldberg, MD, Warren Alpert Medical School of Brown University, Providence, RI, USA. Hans House, MD, University of Iowa School of Medicine, Iowa City, IA, USA. Stacey House, MD, Washington University School of Medicine, St. Louise, MO, USA. Timothy Jang, MD, David Geffen School of Medicine at UCLA, Los Angeles, CA, USA. Christopher Kabrhel, MD, Massachusetts General Hospital, Boston, MA, USA. Stephen Lim, MD, Louisiana State University School of Medicine, New Orleans, New Orleans, LA, USA. Troy Madsen, MD, University of Utah School of Medicine, Salt Lake City, UT, USA. Danielle McCarthy, MD, Northwestern University, Feinberg School of Medicine, Chicago, IL, USA. Andrew Meltzer, MD, George Washington University School of Medicine, Washington, DC, USA. Stephen Moore, MD, Penn State Milton S. Hershey Medical Center, Hershey, PA, USA. Mark B. Mycyk, MD, Cook County Hospitals, Chicago, IL, USA. Craig Newgard, MD, Oregon Health and Science University, Portland, OR, USA. Kristen E. Nordenholz, MD, University of Colorado School of Medicine, Aurora, CO, USA. Justine Pagenhardt, MD, West Virginia University School of Medicine, Morgantown, WV, USA. Ithan Peltan, MD, Intermountain Healthcare, Salt Lake City, UT, USA. Katherine L. Pettit, MD, Indiana University School of Medicine, Indianapolis, IN, USA. Michael Pulia, MD, USA. University of Wisconsin School of Medicine and Public Health, Madison, WI, USA. Michael Puskarich, MD, Hennepin County Medical Center and the University of Minnesota, Minneapolis, MN, USA. Lauren Southerland, MD, Ohio State University Medical Center, Columbus, OH, USA. Scott Sparks, MD, Riverside Regional Medical Center, Newport News, VA, USA. Danielle Turner-Lawrence, MD, Beaumont Health, Royal Oak, MI, USA. Marie Vrablik, MD, University of Washington School of Medicine, Seattle, WA, USA. Alfred Wang, MD, Indiana University School of Medicine, Indianapolis, IN, USA. Anthony Weekes, MD, Atrium Health, Charlotte, NC, USA. Lauren Westafer, MD, Baystate Health, Springfield, MA, USA. John Wilburn, MD, Wayne State University School of Medicine, Detroit, MI, USA.

Contributors PG, MAK, AJG, CB, CC and JK conceived and designed the study, JK obtained research funding, designed and organised the registry. MAK and LM provided statistical advice on study design and analysed the data. AJG, MAK and PG drafted the manuscript, and all authors contributed substantially to its revision. PG takes responsibility for the paper as a whole and acts as guarantor.

Funding This study was funded by Department of Emergency Medicine, Indiana University School of Medicine. No Grant Number.

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting or dissemination plans of this research.

Patient consent for publication Not applicable.

Ethics approval The RECOVER registry protocol was reviewed by the institutional review boards (IRBs) at all sites; three IRBs provided approval with waiver of informed consent. All others provided an exemption from human subjects designation. All data were anonymised prior to analysis (local IRB: 56234). Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement No data are available.

ORCID iDs
Alexandra June Gordon http://orcid.org/0000-0002-1866-8565
Jeffrey Kline http://orcid.org/0000-0001-7190-3109

REFERENCES

1 Lipsitch M, Swerdlow DL, Finelli L. Defining the epidemiology of Covid-19 — studies needed. *N Engl J Med Overseas Ed* 2020;382:1194–6.
2 Kerlin MP, Costa DK, Davis BS, *et al*. Actions taken by US hospitals to prepare for increased demand for intensive care during the first wave of COVID-19: a national survey. *Chest* 2021;160:519–28.
3 Yang X, Yu Y, Xu J, *et al*. Clinical course and outcomes of critically ill patients with SARS-CoV-2 pneumonia in Wuhan, China: a single-centered, retrospective, observational study. *Lancet Respir Med* 2020;8:475–81.
4 Ferrando C, Mellado-Artigas R, Gea A, *et al*. Patient characteristics, clinical course and factors associated to ICU mortality in critically ill patients infected with SARS-CoV-2 in Spain: a prospective, cohort, multicentre study. *Rev Esp Anestesiol Reanim* 2020;67:425–37.
5 Ferrando-Vivas P, Doidge J, Thomas K, *et al*. Prognostic factors for 30-day mortality in critically ill patients with coronavirus disease 2019: an observational cohort study. *Crit Care Med* 2021;49:102–11.
6 Gallo Marin B, Aghagoli G, Lavine K. Predictors of COVID-19 severity: a literature review. *Rev Med Virol* 2021;31:1–10.
7 Grasselli G, Greco M, Zanella A, *et al*. Risk factors associated with mortality among patients with COVID-19 in intensive care units in Lombardy, Italy. *JAMA Intern Med* 2020;180:1345–55.
8 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;180:1081–9.
9 Ruan Q, Yang K, Wang W, *et al*. Clinical predictors of mortality due to COVID-19 based on an analysis of data of 150 patients from Wuhan, China. *Intensive Care Med* 2020;46:846–8.
10 Wynants L, Van Calster B, Collins GS, *et al*. Prediction models for diagnosis and prognosis of covid-19: systematic review and critical appraisal. *BMJ* 2020;369:m1328.
11 Xu J, Yang X, Yang L, *et al*. Clinical course and predictors of 60-day mortality in 239 critically ill patients with COVID-19: a multicenter retrospective study from Wuhan, China. *Crit Care* 2020;24:394.
12 Zhang S, Guo M, Duan L, *et al*. Development and validation of a risk factor-based system to predict short-term survival in adult hospitalized patients with COVID-19: a multicenter, retrospective, cohort study. *Crit Care* 2020;24:438.
13 Zhao Z, Chen A, Hou W, *et al*. Prediction model and risk scores of ICU admission and mortality in COVID-19. *PLoS One* 2020;15:e0236618.
14 Zhou F, Yu T, Du R, *et al*. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *Lancet* 2020;395:1054–62.
15 Levy TJ, Richardson S, Coppa K. A predictive model to estimate survival of hospitalized COVID-19 patients from admission data. *medRxiv* 2020.
16 Brajer N, Cozzi B, Gao M, *et al*. Prospective and external evaluation of a machine learning model to predict in-hospital mortality of adults at time of admission. *JAMA Netw Open* 2020;3:e1920733.
17 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.
18 Kline JA, Pettit KL, Kabrhel C, *et al*. Multicenter registry of United States emergency department patients tested for SARS-CoV-2. *J Am Coll Emerg Physicians Open* 2020;1:1341–8.
19 Niknam KK, Mitarai T, Nesbitt J, et al. SOFA illness severity score adapted to predict in-hospital mortality in emergency critical care patients. *Am J Emerg Med* 2019.
20 StataCorp. Stata Multiple-Imputation reference manual College Station, 2021. Available: https://www.stata.com/manuals/mi.pdf
21 Cummings P. Estimating adjusted risk ratios for matched and unmatched data: an update. *Stata J* 2011;11:290–8.
22 DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 1988;44:837–45.
23 Newman TB, Kohn MA. *Evidence-based diagnosis : an introduction to clinical epidemiology*. Cambridge ; New York: Cambridge University Press, 2019.
24 Kline JA, Camargo CA, Courtney DM, *et al*. Clinical prediction rule for SARS-CoV-2 infection from 116 U.S. emergency departments 2-22-2021. *PLoS One* 2021;16:e0248438.
25 Mackey K, Ayers CK, Kondo KK, *et al*. Racial and Ethnic Disparities in COVID-19-Related Infections, Hospitalizations, and Deaths : A Systematic Review. *Ann Intern Med* 2021;174:362–73.
26 Lowe KE, Zein J, Hatipoğlu U, *et al*. Association of smoking and cumulative Pack-Year exposure with COVID-19 outcomes in the Cleveland clinic COVID-19 registry. *JAMA Intern Med* 2021;181:709.
27 Cho S-Y, Park S-S, Song M-K, *et al*. Prognosis score system to predict survival for COVID-19 cases: a Korean nationwide cohort study. *J Med Internet Res* 2021;23:e26257.
28 Berenguer J, Borobia AM, Ryan P, *et al*. Development and validation of a prediction model for 30-day mortality in hospitalised patients with COVID-19: the COVID-19 SEIMC score. *Thorax* 2021;76:920–9.
29 Nicholson CJ, Wooster L, Sigurslid HH. Estimating risk of mechanical ventilation and mortality among adult COVID-19 patients admitted to mass General Brigham: the vice and dice scores. *medRxiv* 2020.
30 Magro B, Zuccaro V, Novelli L. Predicting in-hospital mortality from coronavirus disease 2019: a simple validated APP for clinical use. *PLoS One* 2021;16:e0245281.
31 King JT, Yoon JS, Rentsch CT. Development and validation of a 30-day mortality index based on pre-existing medical administrative data from 13,323 COVID-19 patients: the Veterans health administration COVID-19 (VACO) index. *PLoS One* 2020;15:e0241825.
32 Goodacre S, Thomas B, Sutton L, *et al*. Derivation and validation of a clinical severity score for acutely ill adults with suspected COVID-19: the Priest observational cohort study. *PLoS One* 2021;16:e0245840.
33 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;9:349–59.
34 Hohl CM, Rosychuk RJ, Archambault PM. Derivation and validation of a clinical score to predict death among non-palliative covid-19 patients presenting to emergency departments: the CCEDRRN COVIDmortality score. *medRxiv* 2021.
35 Halasz G, Sperti M, Villani M, *et al*. A machine learning approach for mortality prediction in COVID-19 pneumonia: development and evaluation of the Piacenza score. *J Med Internet Res* 2021;23:e29058.
36 Gupta RK, Marks M, Samuels THA, *et al*. Systematic evaluation and external validation of 22 prognostic models among hospitalised adults with COVID-19: an observational cohort study. *Eur Respir J* 2020;56:2003498.
37 Update to living systematic review on prediction models for diagnosis and prognosis of covid-19. *BMJ* 2021;372:n236.
38 RECOVERY Collaborative Group, Horby P, Lim WS, *et al*. Dexamethasone in hospitalized patients with Covid-19. *N Engl J Med* 2021;384:693–704.