 doi:10.20944/preprints202102.0059.v1

Development and validation of a prediction model for 28-day in-hospital mortality in critically ill patients with COVID-19

Paloma Ferrando-Vivas (0000-0002-2163-645X), Doug W Gould (0000-0003-4148-3312), James C Doidge (0000-0002-3674-3100), Karen Thomas (0000-0001-7548-4466), Paul Mouncey (0000-0002-8510-8517), Manu Shankar-Hari (0000-0002-5338-2538), J Duncan Young (0000-0002-6838-4835), Kathryn M Rowan (0000-0001-8217-5602), David A Harrison (0000-0002-9002-9098), on behalf of the ICNARC Covid-19 Team†

†Group members listed in Acknowledgments

Intensive Care National Audit & Research Centre (ICNARC), 24 High Holborn, London WC1V 6AZ, United Kingdom

Paloma Ferrando-Vivas, Statistician

Doug W Gould, Senior Researcher

James C Doidge, Senior Statistician

Karen Thomas, Senior Statistician

Paul Mouncey, Head of Research

Kathryn M Rowan, Director

David A Harrison, Head Statistician

Intensive Care Unit, Guy's and St Thomas' NHS Foundation Trust, and School of Immunology & Microbial Sciences, Kings College London, London, United Kingdom

Manu Shankar-Hari, NIHR Clinician Scientist

Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford, UK

J Duncan Young, Professor of Intensive Care Medicine

Correspondence to: Professor David Harrison david.harrison@icnarc.org

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 1 February 2021 doi:10.20944/preprints202102.0059.v1

Abstract

Objectives

To develop and validate a prediction model for 28-day in-hospital mortality among adult patients critically ill with COVID-19 in the UK.

Design

Observational cohort study.

Setting

287 adult critical care units in England, Wales and Northern Ireland, of which 260 admitted at least one eligible patient.

Participants

10,933 patients with confirmed COVID-19 of whom 10,401 were eligible (excluding 532 patients with a duration of critical care less than 24 hours and 1 patient with unknown 28-day outcome): 8,666 development (March-April 2020) and 1,735 temporal validation (May-August 2020).

Main outcome measures

28-day in-hospital mortality from start of critical care.

Results

Two models were developed using 14 patient level predictors selected from 30 candidate predictors, with and without adjustment for calendar time. In the temporal validation data, the model discrimination was maintained (c index 0.78) but calibration was poor, particularly for the model not adjusted for calendar time (ratio of observed to predicted mortality 0.74 versus 0.88 for the model adjusted for calendar time).

Conclusions

We developed and validated a prediction model for 28-day in-hospital mortality for patients critically ill with COVID-19. Although absolute predictions were inaccurate due to changing outcomes, the models will support risk-adjustment in analyses and monitoring changes in risk-adjusted outcome over time.

Key words: COVID-19; Critical Illness; Hospital Mortality; Intensive Care Unit; Outcome Assessment, Health Care; Prognosis; Risk Adjustment; Severe acute respiratory syndrome coronavirus 2

Introduction

Numerous statistical models have been developed for patients with COVID-19, including both diagnostic models to identify likely COVID-19 and prognostic models to predict outcomes including mortality, disease progression, ventilation and length of hospital stay.(1) However, substantial limitations have been identified in the development, validation and reporting of these models, and none of the models to date have focussed on the most severely ill patients, those admitted to critical care units. There is a long history of prediction models in adult critical care. These models take information from early in the patient's critical illness and make a prediction about the patient's likely outcome. These can then be used to support risk-adjusted analyses and to monitor changing outcomes over time, or pooled to compare observed and expected outcomes for critical care providers.(2)

In the UK, the Intensive Care National Audit and Research Centre (ICNARC) co-ordinates the Case Mix Programme. This is the national clinical audit covering all NHS adult, general intensive care and combined intensive/high dependency care units in England, Wales and Northern Ireland, and some additional specialist and non-NHS critical care units, which undertakes risk-adjusted benchmarking of critical care providers. Since the start of the UK epidemic, ICNARC has been collecting data from critical care units/areas participating in the Case Mix Programme to support policy makers and providers of care within the NHS, with rapid data collation and analysis. ICNARC has developed prediction models that underpin the Case Mix Programme.(3) However, the generic ICNARC model performs poorly when evaluated in patients with COVID-19, identifying the need to develop a specific model. Here we present the development and validation of the ICNARC COVID-19 risk prediction model for adult patients with confirmed COVID-19 treated in critical care units.

Methods

Selection of data

For model development, data were extracted from the Case Mix Programme database for all patients in England, Wales and Northern Ireland admitted to critical care between 1 March 2020 and 30 April 2020 with a diagnosis of COVID-19 confirmed either at, or after, the start of critical care. Initial model development was undertaken on a dataset locked for analysis on 21 July 2020. The final model was re-estimated using an updated dataset locked on 8 January 2021. Patients with a duration of critical care of less than 24 hours were excluded, as this was the timeframe over which potential predictors were measured. We also excluded patients transferred to another critical care unit within 28 days of admission who had an unknown final critical care outcome. For patients with multiple critical care admissions, only the first admission was included.

Models were validated using data for patients admitted between 1 May and 31 August 2020 (locked for analysis on 8 January 2021). The same inclusion and exclusion criteria were applied.

Outcome and potential predictors

The primary outcome was in-hospital mortality at 28 days following the start of critical care (28-day in-hospital mortality). Patients discharged to a non-acute setting prior to 28 days were considered to have survived. The selection of 28 days was informed by analysis of longer-term outcomes showing that the majority of hospital deaths in critically ill patients with COVID-19 occurred by 28 days following the start of critical care.(4)

Potential predictors were selected, a priori, based on: established relationships with outcome for critically ill patients; emerging information from the COVID-19 pandemic, including evaluation of prognostic factors using the Case Mix Programme database (5); and availability within the Case Mix Programme dataset (eTable 1).

All physiological and laboratory variables were assessed as the most extreme values within the first 24 hours of critical care. Where patients were transferred or readmitted to critical care, physiological and laboratory variables were included from the first 24 hours of the first admission only. To avoid adjusting out differences in patient outcome related to treatment, variables such as organ support received were not assessed for inclusion in the models. Missing data were handled with multiple imputation (see Supplementary Material for detail).

Model development and validation

Due to outcomes for critically ill patients with COVID-19 improving over time,(6) which may reflect changes in treatment or clinical management, we developed two versions of the model: incorporating and not incorporating calendar time, such that a model with or without adjustment for time can be selected depending on the intended use of the model.

Full details of model development and validation can be found in the Supplementary Material. Briefly, missing data in predictors were imputed to produce 10 multiply imputed datasets. The predictor-outcome relationship for continuous predictors was explored using restricted cubic splines with up to five knots. A parsimonious physiology model was developed by selecting significant predictors ($P<0.1$) from a multivariable model with all physiological predictors. All other predictors were then added to the model and a similar selection process was applied. Predictors were then removed from the model one-by-one and only predictors making an important contribution were retained. Pre-specified interaction terms were added to the model and retained if they made an important contribution. Finally, calendar time was added to the model, and the contribution of all other predictors was reassessed.

The models were validated in both the development and temporal validation cohorts. When validated using the temporal validation cohort, calendar time was set to 30 April. Discrimination was assessed with the c index (area under the receiver operating characteristic curve), accuracy was assessed with Brier score, and calibration was assessed graphically and using Cox calibration regression.

Patient and Public Involvement

We did not directly include PPI in this evaluation, but the database used in the study was developed with PPI involvement and ICNARC is overseen by a Board of Management (Trustees) that includes patient and public representatives.

 doi:10.20944/preprints202102.0059.v1

Results

Model development

ICNARC received data for 11,079 admissions (including transfers between critical care units and readmissions) of 9,080 patients with COVID-19 admitted to 255 of 287 participating critical care units between 1 March 2020 and 30 April 2020. The remaining 32 units all confirmed that no patients were admitted with COVID-19 during this period. After excluding patients with a duration of critical care of less than 24 hours (n=413, 4.5%) and one patient with unknown outcome, a total of 8,666 patients were included in model development (Figure 1). Patient characteristics are presented in Table 1 and eTable 4. There were 3,464 (40.0%) in-hospital deaths by 28 days following the start of critical care.

The model development steps are summarised in eTable 5 and full model coefficients are provided in eTable 6. The final models included the following predictors: age; quintile of deprivation; ethnicity; body mass index; any dependency prior to hospital admission; any severe condition in the past medical history; highest heart rate; highest respiratory rate; PaO_2/FiO_2 ratio from the arterial blood gas with the lowest PaO_2; highest blood lactate concentration; highest serum creatinine concentration; highest serum urea concentration; neutrophil count associated with the lowest white blood count; and lowest platelet count.

Model validation

ICNARC received data for 2,521 admissions of 1,853 patients with COVID-19 admitted to 229 of the 287 participating critical care units between 1 May 2020 and 31 August 2020 (again, all remaining units confirmed that no patients were admitted with COVID-19). After excluding patients with a duration of critical care of less than 24 hours (n=118, 6.3%), a total of 1,735 patients were included in the temporal validation (Figure 1, Table 1 and eTable 4). Overall, 28-day in-hospital mortality in the validation cohort was substantially lower than in the development cohort (30.3% vs 40.0%).

Model performance in the development and validation cohorts is summarised in Table 2 and calibration plots are presented in Figure 2. Discrimination was well maintained in the temporal validation, with a c index of 0.78. However, both models

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 1 February 2021 doi:10.20944/preprints202102.0059.v1

overpredicted mortality in the validation cohort. Unsurprisingly, this discrepancy was greater for the model not incorporating calendar time.

Discussion

We have developed prediction models for 28-day in-hospital mortality among critically ill patients with COVID-19, with and without adjustment for time trend. The ability of the models to discriminate between survivors and non-survivors remained similar when the models were validated using data from a later time period, but the models were poorly calibrated, reflecting improvements in outcomes over the time period studied.

The major strength of this work is the large, national database, with 100% coverage of general critical care units providing Level 3 (intensive) care. Harnessing ongoing, routine data collection meant that established systems were already in place with trained data collectors following existing definitions. The disadvantage of this approach was that only routinely recorded fields were available, meaning that some variables that have been found to be useful predictors of outcomes for hospitalised patients with COVID-19 in other studies, for example C reactive protein and lymphocyte count,(1, 7) were not available for inclusion in this model. Despite this, the model demonstrated discrimination comparable with some of the most extensively validated models for hospitalised patients.(1, 8, 9)

The poor calibration in the temporal validation data was to be expected given previously reported improving outcomes for critically ill patients with COVID-19 over the course of the first epidemic wave.(6) While this limits the scope for applying the models to prediction of future patient outcomes, the models do provide baseline predictions to support monitoring of changes in risk-adjusted outcome over time. Nesting the models within ongoing data collection will allow further recalibration and further development to take place. Our focus was a service evaluation of critical care in the UK, and the models require evaluation and recalibration for use in other settings.

Acknowledgements

We thank all those working in critical care units across England, Wales and Northern Ireland and contributing to the care of patients, and, particularly, those responsible for submitting data rapidly and regularly during the COVID-19 epidemic.

Additional ICNARC Covid-19 Team members: Yemi Banjo, Kasia Borowczak, Tom Cousins, Peter Cummins, Keji Dalemo, Robert Darnell, Hanna Demissie, Laura Drikite, Andrew Fleming, Ditte Frederiksen, Sarah Furnell, Abdo Hussein, Abby Koelewyn, Tim Matthews, Izabella Orzechowska, Sam Peters, Alvin Richards-Belle, Tyrone Samuels, Michelle Saull.

Footnotes

Contributors: Conceptualisation and design: PFV, DWG, JCD, KT, PM, MSH, JDY, KMR, DAH. Analysis: PFV, JCD, KT, DAH. Writing original draft: PFV, DWG, DAH. Reviewing and editing: JCD, KT, PM, MSH, JDY, KMR, DAH. Supervision: KMR, DAH. DAH is guarantor and corresponding author for this work, and attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

Funding: No specific funding received. MSH is supported by the National Institute for Health Research Clinician Scientist Award (CS-2016-16-011). The views expressed in this publication are those of the author(s) and not necessarily those of the NHS, the National Institute for Health Research or the Department of Health and Social Care.

Competing interests: All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare: no support from any organisation for the submitted work; no financial relationships with any organisation that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

Ethical approval: The Case Mix Programme has approval for the collection and use of patient-identifiable data without consent under Section 251 of the NHS Act 2006 (approval number PIAG 2–10(f)/2005). As a service evaluation, approval from a Research Ethics Committee was not required (UK Health Research Authority).

Data sharing: Requests for accessing data from the Case Mix Programme are subject to approval by an independent Data Access Advisory Group (see https://www.icnarc.org/Our-Audit/Cmp/Reports/Access-Our-Data for more details). Requests should be submitted to the corresponding author in the first instance.

Dissemination declaration: It is not possible to disseminate results directly to study participants. Information on all publications arising from the Case Mix Programme is made available to the public on the ICNARC website at https://www.icnarc.org/Publications.

Transparency declaration. The corresponding author (the manuscript's guarantor) affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned (and, if relevant, registered) have been explained.

References

1. Wynants L, Van Calster B, Collins GS, Riley RD, et al: Prediction models for diagnosis and prognosis of covid-19 infection: systematic review and critical appraisal. *BMJ* 2020; 369:m1328
2. Power GS, Harrison DA: Why try to predict ICU outcomes? *Curr Opin Crit Care* 2014; 20(5):544-549
3. Ferrando-Vivas P, Jones A, Rowan KM, Harrison DA: Development and validation of the new ICNARC model for prediction of acute hospital mortality in adult critical care. *J Crit Care* 2017; 38:335-339
4. Richards-Belle A, Orchezowska I, Doidge J, Thomas K, et al: Critical care outcomes, for the first 200 patients with confirmed COVID-19, in England, Wales and Northern Ireland: a report from the ICNARC Case Mix Programme. *J Intensive Care Soc* 2020; In press
5. Ferrando-Vivas P, Doidge J, Thomas K, Gould DW, et al: Prognostic factors for 30-day mortality in critically ill patients with coronavirus disease 2019: an observational cohort study. *Crit Care Med* 2020; In press
6. Doidge JC, Mouncey PR, Thomas K, Gould DW, et al: Trends in intensive care for patients with COVID-19 in England, Wales and Northern Ireland. *Preprints* 2020:2020080267
7. Henry BM, de Oliveira MHS, Benoit S, Plebani M, et al: Hematologic, biochemical and immune biomarker abnormalities associated with severe illness and mortality in coronavirus disease 2019 (COVID-19): a meta-analysis. *Clin Chem Lab Med* 2020; 58(7):1021-1028
8. Gupta RK, Marks M, Samuels THA, Luitel A, et al: Systematic evaluation and external validation of 22 prognostic models among hospitalised adults with COVID-19: an observational cohort study. *medRxiv [preprint]* 2020; doi:10.1101/2020.07.24.20149815
9. Knight SR, Ho A, Pius R, Buchan I, 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

Table 1. Characteristics of the development and validation samples

Characteristics and outcome	Development (N=8,666)	Validation (N=1,735)	All patients (N=10,401)
Demographics			
Age (years), median (IQR)	60 (51, 68)	60 (50, 69)	60 (51, 68)
Sex, n (%)			
Female	2,485 (28.7)	615 (35.4)	3,100 (29.8)
Male	6,176 (71.3)	1,120 (64.6)	7,296 (70.2)
Ethnicity, n (%)			
White	5,426 (62.6)	1,166 (67.2)	6,592 (63.4)
Mixed	165 (1.9)	19 (1.1)	184 (1.8)
Asian	1,314 (15.2)	295 (17.0)	1,609 (15.5)
Black	872 (10.1)	87 (5.0)	959 (9.2)
Other	602 (6.9)	76 (4.4)	678 (6.5)
Not stated	287 (3.3)	92 (5.3)	379 (3.6)
Quintile of deprivation *, n (%)			
1 (least deprived)	1,260 (14.7)	189 (11.1)	1,449 (14.1)
2	1,399 (16.4)	253 (14.8)	1,652 (16.1)
3	1,680 (19.6)	299 (17.5)	1,979 (19.3)
4	2,104 (24.6)	386 (22.6)	2,490 (24.3)
5 (most deprived)	2,113 (24.7)	579 (33.9)	2,692 (26.2)
Body mass index (kg m^{-2}), median (IQR)	28.4 (25.0, 32.9)	28.4 (24.6, 33.8)	28.4 (24.9, 33.0)
Medical history			
Dependency prior to hospital admission, n (%)			
Able to live without assistance in daily activities	7,830 (91.2)	1,437 (83.4)	9,267 (89.9)
Some (minor/major) assistance with daily activities	734 (8.5)	273 (15.9)	1,007 (9.8)
Total assistance with all daily activities	23 (0.3)	12 (0.7)	35 (0.3)
Severe conditions in the past medical history †, n (%)			
Respiratory disease	80 (0.9)	35 (2.0)	115 (1.1)
Cardiovascular disease	42 (0.5)	20 (1.2)	62 (0.6)
Liver disease	23 (0.3)	23 (1.3)	46 (0.4)
Renal disease	134 (1.6)	34 (2.0)	168 (1.6)
Immunocompromise	374 (4.4)	112 (6.5)	486 (4.7)
Acute severity scores from first 24h of critical care, median (IQR)			
APACHE II score	14 (11, 18)	15 (12, 19)	14 (11, 18)
ICNARC physiology score	20 (15, 25)	16 (12, 21)	19 (14, 25)
Interventions during first 24h of critical care, n (%)			
Sedated for entire of first 24h	4,071 (47.3)	327 (18.9)	4,398 (42.5)
Mechanical ventilation at any time in first 24h	5,518 (64.4)	603 (35.0)	6,121 (59.5)
Outcome			
28-day outcome ‡, n (%)			
Alive	5,202 (60.0)	1,210 (69.7)	6,412 (61.6)
Dead	3,464 (40.0)	525 (30.3)	3,989 (38.4)

APACHE, Acute Physiology And Chronic Health Evaluation; ICNARC, Intensive Care National Audit & Research Centre; IQR, interquartile range.

* Based on the patient's usual residential postcode using English Index of Multiple Deprivation 2019, Welsh Index of Multiple Deprivation 2019 or Northern Irish Multiple Deprivation Measure 2017

† Defined according to APACHE II (see Supplementary Material).

‡ Patients discharged to a non-acute setting within 28 days are assumed to survive to 28 days.

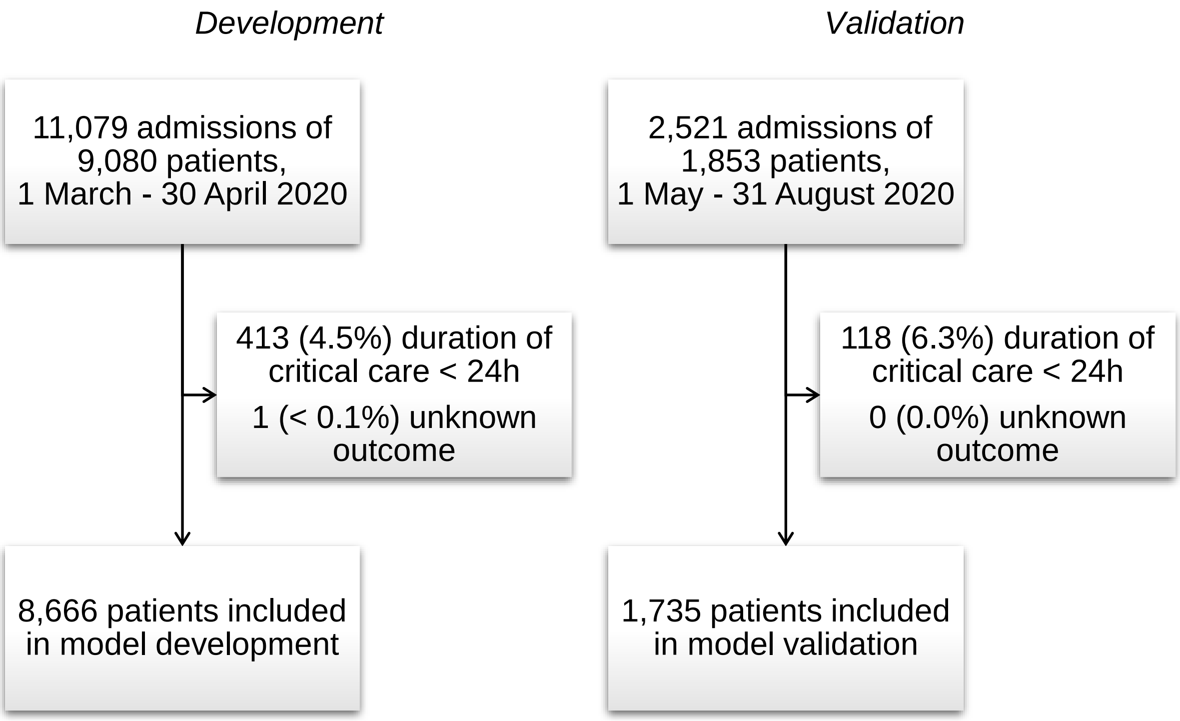
Table 2. Model performance

Model	Model not incorporating calendar time	Model incorporating calendar time
Internal validation (N=8,666)		
c index (95% CI) [% optimism]	0.770 (0.760, 0.780) [0.5%]	0.772 (0.762, 0.782) [0.6%]
Brier score (R-squared)	0.189 (0.222)	0.188 (0.226)
Temporal validation (N=1,735)		
c index (95% CI)	0.780 (0.757, 0.802)	0.781 (0.756, 0.803)
Brier score (R-squared)	0.180 (0.149)	0.168 (0.202)
Observed/predicted mortality (95% CI)	0.74 (0.68, 0.79)	0.88 (0.81, 0.94)
Cox calibration regression		
Intercept (95% CI)	−0.61 (−0.73, −0.49)	−0.25 (−0.37, −0.12)
Slope (95% CI)	0.99 (0.88, 1.11)	1.00 (0.89, 1.12)

CI, confidence interval.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 1 February 2021 doi:10.20944/preprints202102.0059.v1

Figure 1. Flow diagram of study inclusion



Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 1 February 2021 doi:10.20944/preprints202102.0059.v1

Figure 2. Model calibration: (A) internal validation, not incorporating calendar time; (B) internal validation, incorporating calendar time; (C) temporal validation, not incorporating calendar time; (D) temporal validation, incorporating calendar time

The points show the observed 28-day in-hospital mortality plotted against the predicted mortality in ten equal sized groups by predicted mortality; the vertical lines through each point are 95% confidence intervals on the observed mortality; the dashed lines show locally-weighted scatterplot smoothing (LOWESS) applied to the observed and predicted log odds of mortality; the grey diagonal line indicates perfect calibration.