Evaluating risk stratification scoring systems to predict mortality in patients with COVID-19

Kelly Chu ![ORCID] , Batool Alharahsheh, Naveen Garg, Payal Guha

Faculty of Medicine, Imperial College London, London, UK

Correspondence to
Kelly Chu; kc1420@ic.ac.uk

ABSTRACT

Background The COVID-19 pandemic has necessitated efficient and accurate triaging of patients for more effective allocation of resources and treatment.

Objectives The objectives are to investigate parameters and risk stratification tools that can be applied to predict mortality within 90 days of hospital admission in patients with COVID-19.

Methods A literature search of original studies assessing systems and parameters predicting mortality of patients with COVID-19 was conducted using MEDLINE and EMBASE.

Results 589 titles were screened, and 76 studies were found investigating the prognostic ability of 16 existing scoring systems (area under the receiving operator curve (AUROC) range: 0.550–0.966), 38 newly developed COVID-19-specific prognostic systems (AUROC range: 0.6400–0.9940), 15 artificial intelligence (AI) models (AUROC range: 0.840–0.955) and 16 studies on novel blood parameters and imaging.

Discussion Current scoring systems generally underestimate mortality, with the highest AUROC values found for APACHE II and the lowest for SMART-COP. Systems featuring heavier weighting on respiratory parameters were more predictive than those assessing other systems. Cardiac biomarkers and CT chest scans were the most commonly studied novel parameters and were independently associated with mortality, suggesting potential for implementation into model development. All types of AI modelling systems showed high abilities to predict mortality, although none had notably higher AUROC values than COVID-19-specific prediction models. All models were found to have bias, including lack of prospective studies, small sample sizes, single-centre data collection and lack of external validation.

Conclusion The single parameters established within this review would be useful to look at in future prognostic models in terms of the predictive capacity their combined effect may harness.

INTRODUCTION

The SARS-CoV-2 outbreak has put enormous strain on healthcare systems around the world. According to the WHO, as of 12 January 2021, there have been more than 91 million cases of COVID-19 reported worldwide, with almost 2 million deaths.[1] To properly allocate resources and aid clinical decision-making, there is an urgent need for a simple, accurate system to rapidly identify patients who are at the highest risk of death.

Traditionally, scoring systems are used in healthcare to stratify risk, predict outcomes and appropriately manage patients.[2] For example, the CRB-65 scoring system is efficiently used to assess the mortality risk of pneumonia in primary care to determine the need for management escalation.[3]

Risk stratification methods have been effectively used in previous viral outbreaks such as the Ebola epidemic in 2014 to reduce casualties.[4] With COVID-19 being a novel disease, no pre-existing risk stratification methods were available, so traditional scoring systems were adapted in the early stages of the pandemic. As the pandemic progressed, COVID-19-specific tools were developed by studying patients' characteristics relating strongly to mortality and incorporating them into scoring systems.

Although artificial intelligence (AI) algorithm development varies depending on the number of possible outcomes, it is an ideal way of stratifying patients.[5] It uses dynamic data and continual updating of its algorithm to increase the accuracy of its predictions.

This review aims to provide a summary of the literature available on risk stratification tools, including prediction models and single parameters used to predict the mortality of patients with COVID-19 to aid clinical decision-making. This review also aims to evaluate the applications of AI in mortality prediction models.

This study hopes to fill in the gaps in the current literature reviewing human and AI scoring tools. In addition, new studies investigating parameters associated with SARS-CoV-2 mortality are being published; therefore, constant evaluation of risk stratification tools is imperative in a rapidly evolving pandemic.

COVID-19 (TI, AB, KW)	Risk stratification (TI, AB, KW)	Mortality (TI, AB, KW)
COVID-19	Prognos*4 adj2 model/score/algorithm /tool	Hospital Mortality (MeSH)
COVID-2019	Clinical decision tool	Death*1
SARS-CoV-2	Predicti* adj2 model/score/algorithm /tool	Mortality
Severe acute respiratory syndrome coronarvirus 2	Risk adj2 model / predicti*/score/tool/stratification	Fatal*5
2019-nCoV	Scor*3 system*1	
	Mortality adj1 scor*3	

The following search concepts were combined using Boolean operators: COVID-19 (TI, AB, KW) AND Risk stratification (TI, AB, KW) AND Mortality (TI, AB, KW)
AB, abstract; KW, keywords; TI, title, the '/' indicated a different variation.

METHODS

A comprehensive search of MEDLINE and EMBASE between 1 January 2019 and 5 January 2021 was conducted to retrieve studies related to mortality risk prediction of patients with COVID-19. The search was done using the keywords and relevant MeSH terms displayed in table 1.

Inclusion criteria were the following: (1) primary studies carried out on adult patients who are COVID-19-positive; (2) reporting of a model for predicting mortality with a reported area under the receiving operator curve (AUROC) value; and (3) routine blood or imaging parameters with mortality as the main outcome of interest. The established definition of AUROC applied to the context of a COVID-19 mortality prediction model was used; the accuracy of the model was used to discriminate the mortality risk levels in patients with COVID-19 .[6]

Exclusion criteria were non-English studies, sample size <100 patients and non-peer-reviewed publications. Any disagreements during screening were resolved by consensus. Mortality, for this review, is defined as death within 90 days of hospital admission due to COVID-19.

A data extraction form was generated to synthesise the following information: study title, method of calculation of the model or examined parameters (eg, statistical modelling or analysis, AI), scoring system versus analysis of single parameters, 'summary of included parameters and AUROC for scoring systems', 'name and category of parameter (eg, biomarker)' for single parameters and any additional salient findings.

RESULTS

After deduplication of original search results, title and abstracts of 589 studies were screened for relevance, and subsequently full-text articles were obtained and further assessed for eligibility. In all, 76 studies were identified that would inform our review.

Adapted current scoring systems

The sudden arrival of the pandemic has necessitated the application of existing prognostic systems to triage the influx of patients with COVID-19 to optimise distribution of limited resources and treatment. The accuracy of scoring systems adapted for COVID-19 mortality is detailed in online supplemental table 1 and then analysed to explore potential reasons for their differing predictive ability of mortality in patients with COVID-19 .

Scoring systems are listed in order of descending AUROC values, as methodical differences between studies deem it inappropriate to merge AUROC results. For example, the *Quick Sequential Organ Function Assessment* (qSOFA) AUROC values ranged from 0.6200 to 0.8860 (online supplemental table 1), possibly due to different cut-off points. In addition, mortality was measured by 72 hours in some studies and up to 90 days in others, and sample sizes ranged from 105 to 864 across studies (online supplemental table 1).

The *Acute Physiology and Chronic Health Evaluation II* (APACHE II) score was found to have the highest AUROC values, followed by *Modified Elixhauser Index* (mEI) and *Sequential Organ Function Assessment* (SOFA) systems. APACHE II presides over other scores in terms of mortality prediction possibly due to its consideration of both age and comorbidities, whereas scores such as CURB-65 only assesses age and SOFA involves neither. Notably, however, the cut-off value for APACHE II is much lower when applied to patients with COVID-19 than under normal intensive care unit (ICU) conditions; while *Glasgow Coma Scale* (GCS) is an important component of APACHE II, the nervous system is typically less impacted than the respiratory system in COVID-19 infection.[7]

COVID-19 scoring systems

Prediction scores play a vital role in guiding clinical decision-making for hospitalised patients with COVID-19. Online supplemental table 2 summarises recently developed scores and their AUROC values.

Different risk stratification tools use a variety of parameters to predict mortality. Online supplemental table 3 summarises the most common parameters used in novel COVID-19 mortality prediction scores. The two parameters associated with high predictive performance (higher AUROC) were lymphocyte count and D-dimer, with age being the most consistently used parameter. The most common parameter used in novel prediction models for mortality of patients with COVID-19 is age, followed by lymphocyte count, D-dimer, oxygen saturation, C reactive protein (CRP) and platelet count. Other less common parameters include respiratory rate (RR), lactate dehydrogenase, neutrophil-to-lymphocyte ratio (NLR), procalcitonin (PCT) and blood urea nitrogen.

The most common comorbidities for predicting mortality are hypertension (HTN), diabetes mellitus

(DM), obesity, cardiovascular disease, chronic kidney disease, smoking and malignancy.

Single parameters

COVID-19 has a different clinical picture to pneumonia and influenza, providing an avenue to explore what routinely available clinical information best predicts mortality. We explored blood parameters and imaging not currently extensively implemented into existing COVID-19 mortality prediction models, which are represented in online supplemental table 4.

Studies examining the associations of a range of laboratory biochemical tests and imaging at admission with mortality for patients with COVID-19 are extensive in the literature. Continued rapid identification of biomarkers that can accurately predict the likelihood of mortality is essential and has been proposed, including inflammatory, coagulation, renal, liver and cardiac biomarkers (online supplemental table 4).

Imaging, particularly chest CT scans, has been studied, with all three studies reporting independent associations with mortality, shown in online supplemental table 5. Alongside prognostic scores developed to assess risk of death, these must be updated to reflect the identification of imaging modalities that may need to be added or replace parameters in existing scores.

AI in predicting mortality

Machine learning (ML) is a subset of AI allowing systems to automatically improve based on new experiences.[8] Online supplemental table 6 illustrates an overview of studies that used ML to predict mortality in patients with COVID-19.

Papers that used ML models have an AUROC greater than 0.8, conveying good discrimination of patients with high mortality risk.[6]

Models with a greater number of incorporated parameters did not find improvements in AUROC score. One model by Yuan et al[9] had a high AUROC of 0.9551 when looking at three parameters, while the model by Vaid et al[10] had a lower AUROC of 0.8400 when looking at 73 different parameters. This suggests that the total number of parameters was a less important factor than the interaction between the parameters in predicting mortality.

Deep learning (DL) is a subset of ML which uses algorithms to analyse multiple factors simultaneously[11]; therefore, it would be more appropriate to handle multiple parameters. Online supplemental table 7 illustrates an overview of the studies that used ML to predict mortality in patients with COVID-19.

There are fewer studies assessing DL models, but similar to ML, these studies possess an AUROC >0.8.

DISCUSSION
Adapted current scoring systems

The variables used within existing scoring systems featured in online supplemental table 1 were analysed to explore potential reasons for their differing predictive ability of mortality in patients with COVID-19.

The APACHE II score was found to have the highest AUROC values, followed by mEI and SOFA systems. APACHE II presides over other scores in terms of mortality prediction possibly due to its consideration of both age and comorbidities, whereas scores such as CURB-65 only assesses age and SOFA involves neither. Notably, however, the cut-off value for APACHE II is much lower when applied to patients with COVID-19 than under normal ICU conditions; while GCS is an important component of APACHE II, the nervous system is typically less impacted than the respiratory system in COVID-19 infection.[7]

Considering the effects of COVID-19 on respiratory function are more marked than its cardiovascular impacts,[12] it is unsurprising that most of the studies listed in online supplemental table 1 show respiratory parameters such as RR in CURB-65 to be independently more indicative of mortality than blood pressure and confusion, which are more related to haemodynamics. qSOFA's focus on blood pressure and mental state may explain its lower AUROC and poorer predictive performance. Cetinkal et al,[13] however, argue that as previous studies reveal worse clinical outcomes in patients with cardiac injury, non-respiratory variables in the CHA_2D_2VASc system such as older age, DM, HTN and previous cardiovascular disease are valuable parameters for mortality risk stratification. However, AUROC values found for CHA_2D_2VASc remain at the low end compared with other existing scoring systems, despite modifications catered to COVID-19 added to form the m-CHA_2D_2VASc scale. Even this version, with an AUROC higher by 0.06, offers predictive ability similar to univariate NLR and inferior to troponin increase.

Ortiz et al[12] demonstrated A-DROP, a modified version of CURB-65, to provide more accurate mortality prediction than *Pneumonia Severity Index* (PSI), CURB-65, CRB-65, SMART-COP, qSOFA and *National Early Warning Score 2* (NEWS2). Its superior discrimination may be due to its more accurate respiratory function evaluation (oxygen saturation [SpO_2] >90% / arterial oxygen tension [PaO2] <60 mm Hg in A-DROP vs respiratory rate ≥30/min in CURB-65). The modified age cut-off (male >70 / female >75 in A-DROP vs age >65 in CURB-65) may also contribute to A-DROP's advantage when applied to COVID-19, considering the median age of COVID-19 non-survivors is 69 years.[14]

Ultimately, although APACHE II, SOFA, PSI and CURB-65 are well-founded in clinical practice, their requirement for sophisticated patient information makes rapid assessment impossible, an important benefit for triaging patients with COVID-19 in often overrun hospitals. Wang et al's study[7] on MEWS suggests this system can overcome the issue of efficiency as a simple and rapid assessment able to be performed within minutes of patient admission while maintaining equal predictive ability.

Intriguingly, Gupta *et al*[15] evaluated 22 prognostic models (including aforementioned systems), concluding that they should not be recommended for routine clinical implementation because none of them offered incremental value compared with univariable predictors to risk stratify COVID-19 mortality, of which patient's age is a strong predictor of mortality. Similarly, Bradley *et al*[16] concluded that CURB-65, NEWS2 and qSOFA all underestimate the mortality of patients with COVID-19.

COVID-19 scoring systems

To maximise the accuracy and effectiveness of mortality prediction models, novel scores should focus on identifying features that are COVID-19-specific. Examples of complications that are highly associated with COVID-19 include hypercoagulability and inflammation.[17 18] However, only 27% of new prognostic scores included in this review incorporated CRP—an important inflammatory marker. Similarly, thrombopenia has been associated with higher rates of mortality,[19] which reflects the importance of including platelet count in prognostic models, but only 16% of new scores took this into account.

Interestingly, the three prediction models with the highest AUROC values have all used D-dimer and lymphocyte count to predict mortality. This could reflect the importance of these two parameters in COVID-19 pathophysiology. However, these are all single-centre studies tested on significantly smaller sample sizes compared with other models with lower AUROC values. Models tested on a larger population, for instance, Mancilla-Galindo *et al*'s[18] national cohort study with a sample size of 83 779 (AUROC=0.8000), could be more representative and generalisable.

The most common parameter used in novel prediction models for mortality of patients with COVID-19 is age, followed by lymphocyte count, D-dimer, oxygen saturation, CRP and platelet count. Other less common parameters include RR, lactate dehydrogenase, NLR, PCT and blood urea nitrogen.

Fumagalli *et al*[19] report age as the strongest predictor of severe outcomes and mortality. Similarly, Mei *et al*'s[20 21] prognostic model included age as one of five indicators of mortality and reports a strong association between advanced age and death from COVID-19.

There seems to be no association between the number of parameters and the prognostic power and accuracy of a scoring system. Several mortality prediction models with a small number of parameters have had higher AUROC values, for example, Liu *et al*[22] had an AUROC value of 0.9940 with only three variables compared with Mancilla-Galindo *et al*[18] (COVID-GRAM) with an AUROC value of 0.7750 and 10 parameters.

The most common comorbidities for predicting mortality are HTN, DM, obesity, cardiovascular disease, chronic kidney disease, smoking and malignancy.

Single parameters

COVID-19 has a different clinical picture to pneumonia and influenza, providing an avenue to explore what routinely available clinical information best predicts mortality. We explored blood parameters not currently extensively implemented into existing COVID-19 mortality prediction models, which are represented in online supplemental table 4.

We discuss the feasibility of introducing the below blood tests and imaging modalities into routine practice for risk stratification of patients with COVID-19.

Cardiac biomarkers

Cardiac biomarkers were the the most common parameters studied in our literature search. High-sensitivity cardiac troponins have been shown to be independently associated with all-cause mortality in patients with COVID-19 (p<0.05), after accounting for age, sex and comorbidities, shown in online supplemental table 4. High-sensitivity cardiac troponins (hs-cTnI and hs-TnT) are markers of myocardial injury that are currently primarily used in the prognostication of acute coronary syndrome. Despite evidence that 50% with confirmed COVID-19 have elevated cardiac biomarkers at the time of hospital admission, the patient sample sizes are limited in current studies to less than 500 patients and single centres.[22] Cao *et al*[23] retrospectively observed 244 patients and incorporated hs-cTnI into a model of empirical prognostic factors. A proposed cut-off (>20 ng/L serum hs-cTnI levels) yielded an AUROC increase from 0.65 to 0.71 (p<0.01) and demonstrated feasibility of this parameter to increase predictive performance.[24]

Inflammatory biomarkers

Liu *et al*[25] confirmed the independent association of PCT with mortality in a cohort of 1525 patients through retrospective analysis. Due to the large cohort and continued follow-up of PCT levels throughout hospital stay, this study provides stronger evidence for the inclusion of PCT into scoring systems, which has begun to be implemented but is still in the minority of included parameters. Fois *et al*[26] used the same study design and identified the systemic inflammation index (SII) as an independent predictor of mortality. However, the study quality was poor—with only 119 patients and the large number of different inflammation indexes being studied in different combinations. It is unclear whether any clinical utility is offered by implementation of SSI, considering deranged lymphocyte count is already widely established as a useful predictor.[20]

Renal and hepatic function biomarkers

Esposito *et al*[27] identified estimated glomerular filtration rate (using a baseline of 60 mL/min/1.73 m²), and Fu *et al*[24] identified cholestasis and hypoproteinaemia as independent predictors of mortality. Interestingly, as with cardiac biomarkers, these were predictors even after accounting for pre-existing comorbidities. The obvious benefit to clinical practice of renal and hepatic function markers is that they

are routinely done on hospital admission and straightforward to clinicians to score in a system. Replication of large-scale multicentre studies is needed before determining the diagnostic validity of such parameters in the stratification of patients with COVID-19 in a statistical or AI model. It must be acknowledged that additional parameters must be externally validated to determine AUROC values and appropriate cut-offs for parameters.[6]

Lung imaging

Trabulus et al,[28] Francone et al[29] and Xu et al[30] examined the relationship between chest CT findings and mortality, with all three studies reporting independent associations with mortality (p<0.05). Two studies[31][32] used a methodology involving an overall severity score of each scan and proposed defined cut-offs above which there was yield of best predictive value. These cut-offs are of value for clinicians to allocate scans with a high/medium/low rating which can be used to triage patients with COVID-19. However, both these studies have limitations in their methodology and design, which need to be addressed before implementation of CT severity into scoring systems. In the study by Gao et al,[31] follow-up was limited to 24 days; a minimum of at least 28-day mortality is recommended to better reflect the clinical course of COVID-19 in most cases.[7] In addition, both severity score studies were retrospective in nature, which is susceptible to incomplete clinical records and bias in the interpretation of CT by different radiologists. Chest CT while highly sensitive is not a first-line test due to limited resources to CT scan in all COVID-19-positive hospital admissions. Routine implementation of admission CT scans would also carry a radiation burden to patients, which is arguably unnecessary if alternative parameters conferring equal predictive power without additional risk of iatrogenic effects could be used. Perhaps, chest CT is more appropriate in the discharge process of clinically stable, triaged patients with COVID-19 rather than as a first-line test as part of an admission scoring system.

AI in predicting mortality

Between ML and DL models, it is unclear which branch of AI modelling would be superior in predicting mortality due to the similar AUROC values. These similar values can be accounted for by limitations in the study methods.

Within all AI modelling papers, Meng et al[33] and Vaid et al[10] were the only studies that conducted external validation. External validation is an important step to verify the effectiveness of the model in patient population. Internal validation would use the same cohort to test the model, which can lead to overfitting and an inaccurately high AUROC. The models created by Bertsimas et al[34] Gao et al[31] and Meng et al[35] gathered training set data from multiple centres, whereas the other models used single-centre data. Therefore, these models would increase applicability to the general population.

As COVID-19 has only been prevalent for a year, not many models have had the chance to be prospectively tested. Vaid et al[10] produced the only model that was prospectively tested. This is important as it demonstrates the model's real-world performance. Many models with a large number of incorporated parameters included patients with missing values, leading to estimation. This may be useful in clinical practice as not all patients have every test carried out.

It is important to recognise that COVID-19 management and treatment guidelines are constantly being updated, which influences mortality rates. As AI models use dynamic data,[10] reporting of model AUROC in earlier stages of the pandemic may not have been as accurate.

Limitations

There are inherent limitations to this review. Most studies included were single centre and retrospective, whereas multicentre, prospective research may provide more insight. Although AUROC scores are universally accepted outcome measures of the accuracy of prediction models,[6] they are limited in their clinical interpretability as they lack a direct link to individual patient outcomes. Thus, future reviews could use additional performance metrics in addition to AUROC to assess the accuracy of different models.

CONCLUSION

The above systems and parameters have been evaluated for their ability to stratify patients with COVID-19 by mortality risk, with predictive ability depicted as AUROC scores. New scoring systems developed specifically for the pandemic demonstrated higher AUROC scores than currently existing scoring systems adapted for COVID-19. However, the predictive strength of AI systems was not notably higher than pandemic-specific scoring systems, potentially due to time restraints of development and incomplete refining of algorithms. Single parameters extracted from scoring systems, novel biomarkers and imaging modalities were also explored for the ability to predict mortality and potential incorporation into novel risk stratification systems.

As most studies in the current literature were retrospective, we propose further prospective, multicentre studies to validate these variables' diagnostic accuracy and multivariate relationships, which may impact their compounded efficacy for COVID-19 mortality prediction. A meta-analysis would address the limitation of the current review of not being able to directly compare and statistically manipulate AUROC scores found in the literature due to differing cut-off points, study sample sizes and mortality periods used by different studies.

In all, refining strategies to triage patients with COVID-19 can bring immense value to healthcare professionals in their clinical decisions concerning optimal treatment for patients with varying mortality risks and allocating scarce resources effectively.

Competing interests None declared.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available in a public, open access repository.

ORCID iD
Kelly Chu http://orcid.org/0000-0003-0527-373X

REFERENCES

1. World Health Organisation. WHO coronavirus disease (COVID-19) Dashboard.. Available: https://covid19.who.int [Accessed 13 Jan 2021].
2. Oprita B, Aignatoaie B, Gabor-Postole DA. Scores and scales used in emergency medicine. practicability in toxicology. *J Med Life* 2014;7(Spec No. 3):4–7.
3. National Institute for Health and Care Excellence (NICE). Pneumonia in adults [QS110], 2016. Available: https://www.nice.org.uk/guidance/qs110/chapter/Quality-statement-1-Mortality-risk-assessment-in-primary-care-using-CRB65-score [Accessed 13 Jan 2021].
4. Kangbai JB, Heumann C, Hoelscher M, *et al*. Sociodemographic and clinical determinants of in-facility case fatality rate for 938 adult Ebola patients treated at Sierra Leone Ebola treatment center. *BMC Infect Dis* 2020;20:293.
5. Lu H, Li Y, Chen M, *et al*. Brain intelligence: go beyond artificial intelligence. *Mobile Networks and Applications* 2018;23:368–75.
6. Hajian-Tilaki K. Receiver operating characteristic (ROC) curve analysis for medical diagnostic test evaluation. *Caspian J Intern Med* 2013;4:627.
7. Wang L, Lv Q, Zhang X. The utility of MEWS for predicting the mortality in the elderly adults with COVID-19: a retrospective cohort study with comparison to other predictive clinical scores. *PeerJ* 2020;8.
8. Kohavi R, Provost F. Glossary of terms. *Machine Learning* 1998;30:271–4.
9. Yuan Y, Sun C, Tang X, *et al*. Development and validation of a prognostic risk score system for COVID-19 inpatients: a multi-center retrospective study in China. *Engineering* 2020. doi:10.1016/j.eng.2020.10.013. [Epub ahead of print: 28 11 2020].
10. Vaid A, Somani S, Russak AJ, *et al*. Machine learning to predict mortality and critical events in a cohort of patients with COVID-19 in New York City: model development and validation. *J Med Internet Res* 2020;22:e24018.
11. Deng L, Yu D. Deep learning: methods and applications. *FNT in Signal Processing* 2013;7:197–387.
12. Ortiz D, Diaz A, Almasary A, *et al*. 223 evaluation of pneumonia scores in patients hospitalized for COVID-19-Related dyspnea. *Ann Emerg Med* 2020;76:S87.
13. Cetinkal G, Kocas BB, Ser OS, *et al*. Assessment of the modified CHA2DS2VASc risk score in predicting mortality in patients hospitalized with COVID-19. *Am J Cardiol* 2020;135:143–9.
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. 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.
16. Bradley P, Frost F, Tharmaratnam K. The utility of established prognostic scores in COVID-19 hospital admissions: a multi-centre prospective evaluation of CURB-65, NEWS2, and qSOFA. *medRxiv* 2020.
17. Weng Z, Chen Q, Li S, *et al*. ANDC: an early warning score to predict mortality risk for patients with coronavirus disease 2019. *J Transl Med* 2020;18:328.
18. Mancilla-Galindo J, Vera-Zertuche JM, Navarro-Cruz AR, *et al*. Development and validation of the patient history COVID-19 (PH-Covid19) scoring system: a multivariable prediction model of death in Mexican patients with COVID-19. *Epidemiol Infect* 2020;148:e286. doi:10.1017/S0950268820002903
19. Fumagalli C, Rozzini R, Vannini M, *et al*. Clinical risk score to predict in-hospital mortality in COVID-19 patients: a retrospective cohort study. *BMJ Open* 2020;10:e040729.
20. Stefanini GG, Chiarito M, Ferrante G, *et al*. Early detection of elevated cardiac biomarkers to optimise risk stratification in patients with COVID-19. *Heart* 2020;106:1512–8.
21. Mei J, Hu W, Chen Q, *et al*. Development and external validation of a COVID-19 mortality risk prediction algorithm: a multicentre retrospective cohort study. *BMJ Open* 2020;10:e044028.
22. Liu Z-M, Li J-P, Wang S-P, *et al*. Association of procalcitonin levels with the progression and prognosis of hospitalized patients with COVID-19. *Int J Med Sci* 2020;17:2468–76.
23. Cao J, Zheng Y, Luo Z, *et al*. Myocardial injury and COVID-19: serum hs-cTnI level in risk stratification and the prediction of 30-day fatality in COVID-19 patients with no prior cardiovascular disease. *Theranostics* 2020;10:9663–73.
24. Fu L, Fei J, Xu S, *et al*. Liver dysfunction and its association with the risk of death in COVID-19 patients: a prospective cohort study. *J Clin Transl Hepatol* 2020;8:1–9.
25. Liu Q, Song NC, Zheng ZK, *et al*. Laboratory findings and a combined multifactorial approach to predict death in critically ill patients with COVID-19: a retrospective study. *Epidemiol Infect* 2020;148.
26. Fois AG, Paliogiannis P, Scano V, *et al*. The systemic inflammation index on admission predicts in-hospital mortality in COVID-19 patients. *Molecules* 2020;25:5725.
27. Esposito A, Palmisano A, Toselli M. Chest CT–derived pulmonary artery enlargement at the admission predicts overall survival in COVID-19 patients: insight from 1461 consecutive patients in Italy. *Eur Radiol* 2020;23:1.
28. Trabulus S, Karaca C, Balkan II, *et al*. Kidney function on admission predicts in-hospital mortality in COVID-19. *PLoS One* 2020;15:e0238680.
29. Francone M, Iafrate F, Masci GM, *et al*. Chest CT score in COVID-19 patients: correlation with disease severity and short-term prognosis. *Eur Radiol* 2020;30:6808–17.
30. Xu PP, Tian RH, Luo S, *et al*. Risk factors for adverse clinical outcomes with COVID-19 in China: a multicenter, retrospective, observational study. *Theranostics* 2020;10:6372.
31. Gao Y, Cai G-Y, Fang W, *et al*. Machine learning based early warning system enables accurate mortality risk prediction for COVID-19. *Nat Commun* 2020;11:1–10.
32. Bertsimas D, Lukin G, Mingardi L, *et al*. COVID-19 mortality risk assessment: an international multi-center study. *PLoS One* 2020;15:e0243262.
33. Shi Y, Pandita A, Hardesty A. Validation of pneumonia prognostic scores in a statewide cohort of hospitalised patients with COVID-19. *International Journal of Clinical Practice* 2020;24:3–7.
34. Bertsimas D, Lukin G, Mingardi L, *et al*. COVID-19 mortality risk assessment: an international multi-center study. *PLoS One* 2020;15:e0243262.
35. Meng L, Dong D, Li L, *et al*. A deep learning prognosis model help alert for COVID-19 patients at high-risk of death: a multi-center study. *IEEE J Biomed Health Inform* 2020;24:1–3584.