Comparison of severity scores for COVID-19 patients with pneumonia: a retrospective study

To the Editor:

Rapidly progressing hypoxemia and acute respiratory distress syndrome were commonly observed in patients with severe acute respiratory syndrome-coronavirus-2 (SARS-CoV-2) viral pneumonia [1]. Although several severity scores including Pneumonia Severity Index (PSI) [2], CURB-65 and CRB-65 (confusion, (urea >7 mmol·L^{-1}), respiratory rate ⩾30 breaths·min^{-1}, blood pressure <90 mmHg (systolic) ⩽60 mmHg (diastolic), age ⩾65 years), [3], A-DROP [4] and SMART-COP [5] have been developed to identify community acquired pneumonia (CAP) patients at high risk and offer therapeutic advice, the underestimation of risk of death from viral pneumonia in these scores has been reported by previous studies [6, 7]. The National Early Warning Score 2 (NEWS2) was developed by National Health Service (NHS) England [8] and, along with quick sequential organ failure assessment score (qSOFA), was proposed as a candidate for prognostic prediction for severe coronavirus disease 2019 (COVID-19) in the situation of limited medical source [9]. The aim of this study was to compare the accuracy of current score rules in hospitalised patients with COVID-19 pneumonia for predicting the risk of death and evaluate feasibility in improving medical decisions by adopting appropriate scores in clinical practice.

Adult inpatients who were diagnosed as COVID-19 according to World Health Organization (WHO) interim guidance and died/were discharged between 29 December 2019 and 15 February 2020 in Jin Yin-tan Hospital (Wuhan city, China) were retrospectively enrolled in this study. After excluding 689 patients who were still hospitalised as of 15 February 2020, 42 patients with missing key data in their medical records which was essential in scoring, and six deaths within 24 h after admission, we were left with 654 cases, including 521 survivors and 133 non-survivors with intact information to complete calculation of all above scores.

The study was approved by the Research Ethics Commission of Jin Yin-tan Hospital (KY-2020-01.01) and the informed consent was waived by the Ethics Commission.

Information were obtained from electronic medical records. A standardised data collection form (a modified version of the WHO/International Severe Acute Respiratory and Emerging Infection Consortium case record form for severe acute respiratory infections) was used for data extraction. The score on admission of each patient was noted for eight severity score rules, including A-DROP, CURB-65, PSI, SMART-COP, NEWS2, CRB-65 and qSOFA. A-DROP was a modified version of the CURB-65 score rules, including the integrated evaluation of age, dehydration, oxygen saturation measured by pulse oximetry (S_{pO_2}) or arterial oxygen saturation (P_{aO_2}), consciousness and blood pressure [4]. Two researchers were responsible for the accuracy of raw data, and a third party was necessary if doubts existed.

Next-generation sequencing or real-time RT-PCR methods were performed to detect SARS-CoV-2 of respiratory specimens. The PCR re-examination was conducted by throat swab specimens after clinical remission of symptoms. A patient was allowed to be discharged if they achieved clinical improvement and had two throat swab samples negative for SARS-CoV-2 RNA obtained at least 24 h apart.

The illness severity of COVID-19 was defined according to Chinese management guidelines for COVID-19 (version 6.0) [10]. The performance of sensitivity, specificity, or area under the curve (AUC)

@ERSpublications
A-DROP is a reliable tool for risk stratification of death in COVID-19 hospitalised patients on admission https://bit.ly/3iDZipD

Cite this article as: Fan G, Tu C, Zhou F, *et al*. Comparison of severity scores for COVID-19 patients with pneumonia: a retrospective study. *Eur Respir J* 2020; 56: 2002113 [https://doi.org/10.1183/13993003.02113-2020].

was defined as poor with a value <0.5, low with a value between 0.5 and 0.7, moderate with a value between 0.7 and 0.85, and excellent with a value >0.85.

We assessed the predictive performance of A-DROP, CURB-65, PSI, SMART-COP, NEWS2, CRB-65 and qSOFA for in-hospital death by describing receiver operating characteristic curves for each score. Sensitivity, specificity, positive predictive value, negative predictive value and their 95% CIs were calculated. The AUC and 95% CI were estimated to determine the discrimination and net reclassification improvement and integrated discrimination improvement were also estimated to assess the improvement of other scores compared with A-DROP score in death prediction.

A two-sided α <0.05 was considered statistically significant for all statistical tests. Statistical analyses were performed by the SAS software, version 9.4 (SAS Institute Inc.), unless otherwise indicated.

Among all seven scores that were determined by patients' information on admission, A-DROP presented the highest discrimination (AUC 0.87; 95% CI 0.84–0.90), following by CURB-65 (AUC 0.85; 95% CI 0.81–0.89), PSI (AUC 0.85; 95% CI 0.81–0.88), SMART-COP (AUC 0.84; 95% CI 0.80–0.88), NEWS2 (AUC 0.81; 95% CI 0.77–0.85), CRB-65 (AUC 0.80; 95% CI 0.76–0.84), and qSOFA (AUC 0.73; 95% CI 0.69–0.78) in predicting in-hospital death. Taking A-DROP as reference, the AUC contrast showed an insignificant difference between A-DROP and CURB-65 or A-DROP and PSI, while the discrimination of A-DROP was significantly better than any other score rules. Similar differences were also observed with respect to INR and IDI. The positive differences of INR and IDI indicated the discrimination of A-DROP was improved compared with other scores (table 1).

The sensitivity of A-DROP⩾2, PSI⩾3, SMART-COP⩾2, NEWS2⩾5, CRB-65⩾1 and qSOFA⩾1 were moderate, whilst that of CURB-65⩾2 was low for identifying patients at risk of death. The specificity of identifying survivors for CURB-65⩾2 was excellent (0.91, 95% CI 0.89–0.93), followed by A-DROP⩾2, PSI⩾3 and SMART-COP⩾2, whilst the specificity for the rest of the scores were low (table 1).

The accuracy of a variety of severity scores to predict in-hospital death in 654 laboratory confirmed COVID-19 patients admitted to hospital was examined in our study and we found A-DROP was a priority clinical tool for predicting the risk of death for patients with COVID-19 pneumonia, compared with other score systems.

A-DROP, a modified version of CURB-65 [4], showed better accuracy of in-hospital death prediction compared to other current widely used CAP-specific tools. According to previous studies, ARDS was common in severe COVID-19 pneumonia [11, 12]. The rapid progression of diffuse bilateral ground-glass opacities on a computed tomography scan and massive alveolar damage with focal haemorrhage, cellular fibromyxoid exudates and hyaline membrane formation in lung histological examination also suggested a close association between COVID-19 pneumonia and low P_{aO_2}/inspiratory oxygen fraction [13]. The modification of more accurate respiratory function evaluation (S_{pO_2} <90%/P_{aO_2} <60 mmHg in A-DROP *versus* respiratory rate ⩾30 breaths per min in CURB-65) could be one reason for improvement in the discrimination of A-DROP. Another reason may be the modification in age (male >70 years/female >75 years in A-DROP *versus* >65 years in CURB-65). The median age of non-survivors with COVID-19 was reported to be 69 years [14]. Besides, the heavier weight on underlying disease instead of respiratory function in PSI may lead to an underestimated severity of COVID-19 pneumonia, compared with A-DROP.

TABLE 1 The comparison of different clinical prediction rules

Variable	AUC (95% CI)	p-value	Cut-off value	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	NRI	IDI
A-DROP	0.87 (0.84–0.90)	Ref	2	0.80 (0.73–0.87)	0.86 (0.83–0.89)	0.60 (0.52–0.67)	0.94 (0.92–0.96)	Ref	Ref
CURB-65	0.85 (0.81–0.89)	0.2259	2	0.63 (0.55–0.71)	0.91 (0.89–0.93)	0.65 (0.56–0.73)	0.91 (0.88–0.93)	0.12	0.06
PSI	0.85 (0.81–0.88)	0.1876	3	0.77 (0.70–0.84)	0.81 (0.78–0.84)	0.50 (0.44–0.57)	0.93 (0.91–0.96)	0.08	0.07
SMART-COP	0.84 (0.80–0.88)	0.0405	2	0.83 (0.77–0.89)	0.76 (0.72–0.80)	0.46 (0.40–0.53)	0.94 (0.92–0.97)	0.08	0.11
NEWS2	0.81 (0.77–0.85)	0.0045	5	0.79 (0.72–0.86)	0.69 (0.65–0.73)	0.40 (0.34–0.46)	0.93 (0.90–0.95)	0.17	0.16
CRB-65	0.80 (0.76–0.84)	0.0001	1	0.83 (0.77–0.89)	0.69 (0.65–0.73)	0.40 (0.34–0.46)	0.94 (0.92–0.96)	0.15	0.15
qSOFA	0.73 (0.69–0.78)	<0.0001	1	0.82 (0.75–0.89)	0.57 (0.53–0.61)	0.33 (0.28–0.38)	0.93 (0.90–0.95)	0.27	0.24

AUC: area under the curve; PPV: positive predictive value; NPV: negative predictive value; NRI: net reclassification improvement; IDI: integrated discrimination improvement; C(U)RB-65: confusion, (urea >7 mmol·L^{-1}), respiratory rate ⩾30 breaths·min^{-1}, blood pressure <90 mmHg (systolic) ⩽60 mmHg (diastolic), age ⩾65 years; PSI: pneumonia severity index; NEWS2: national early warning score 2; qSOFA: quick sequential organ failure assessment.

NEWS2 score assesses respiration rate, oxygen saturations, systolic blood pressure, heart rate, temperature and level of consciousness, which were easier for use in the emergency department [8]. It proved to be a valid tool for early identification in acutely ill patients with infection [8]. However, without considering the scale of respiratory support therapy, the category of oxygen saturation in NEWS2 score may not reflect the severity of hypoxaemia and lung injury accurately. Lacking markers of other organ dysfunction may be also the reason for its unsatisfying performance.

There are some limitations in the study. First, this is a single-centre study and the intrinsic defects of retrospective studies were unavoidable, for example, scores at different time-points were unavailable, so we could hardly evaluate disease severity dynamically. Secondly, only patients discharged or died were included in this study and those still being hospitalised were excluded. Thirdly, it is unable to evaluate SOFA's performance as results of arterial blood gas tests were absent for most patients in this study.

In summary, A-DROP is a reliable tool for risk stratification of death in COVID-19 hospitalised patients on admission.

Guohui Fan[1,2,7], Chao Tu[3,7], Fei Zhou[2,4,7], Zhibo Liu[2,4,7], Yeming Wang[2,4,5,7], Bin Song[3], Xiaoying Gu[1,2], Yimin Wang[2,4], Yuan Wei[3], Hui Li[2,4], Xudong Wu[3], Jiuyang Xu[6], Shengjin Tu[3], Yi Zhang[2,4], Wenjuan Wu[3,8] and Bin Cao[2,4,5,8]

[1]Institute of Clinical Medical Sciences, China-Japan Friendship Hospital, Beijing, China. [2]Institute of Respiratory Medicine, Chinese Academy of Medical Sciences, National Clinical Research Center for Respiratory Disease, National Center for Respiratory Disease, Beijing, China. [3]Jin Yin-tan Hospital, Wuhan, China. [4]Dept of Pulmonary and Critical Care Medicine, Center of Respiratory Medicine, China-Japan Friendship Hospital, Beijing, China. [5]Dept of Respiratory Medicine, Capital Medical University, Beijing, China. [6]Tsinghua University School of Medicine, Beijing, China. [7]Contributed equally to this work. [8]Wenjuan Wu and Bin Cao contributed equally to this article as lead authors and supervised the work.

Correspondence: Bin Cao, Dept of Pulmonary and Critical Care Medicine, China-Japan Friendship Hospital, Institute of Respiratory Medicine, Chinese Academy of Medical Sciences; National Clinical Research Center for Respiratory Disease, Clinical Center for Pulmonary Infections, Capital Medical University; Tsinghua University-Peking University Joint Center for Life Sciences, No 2, East Yinghua Road, Chaoyang District, Beijing, China. 100029. E-mail: caobin_ben@163.com

Received: 14 April 2020 | Accepted after revision: 29 June 2020

Acknowledgments: We acknowledge all healthcare workers involved in the diagnosis and treatment of patients in Wuhan, China.

Author contributions: B. Cao and W. Wu had the idea for and designed the study, and had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. B. Cao, G. Fan, TC, F. Zhou and Z. Liu drafted the paper. B. Cao, G. Fan, F. Zhou, Z. Liu, Yeming Wang, X. Gu and H. Li performed the analysis, and all authors critically revised the manuscript for important intellectual content and gave final approval for the version to be published. G. Fan, F. Zhou, Z. Liu, B. Song, Yeming Wang, J. Xu, Y. Zhang, S. Tu and X. Wu collected the data. All authors agree to be accountable for the all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Conflict of interest: None declared.

Support statement: This work was supported by the Chinese Academy of Medical Sciences (CAMS) Emergency Project of COVID-19 (2020HY320001). Funding information for this article has been deposited with the Crossref Funder Registry.

References

1 Paules CI, Marston HD, Fauci AS. Coronavirus infections–more than just the common cold. *JAMA* 2020; 323: 707–708.
2 Fine MJ, Auble TE, Yealy DM, *et al.* A prediction rule to identify low-risk patients with community-acquired pneumonia. *N Engl J Med* 1997; 336: 243–250.
3 Lim WS, van der Eerden MM, Laing R, *et al.* Defining community acquired pneumonia severity on presentation to hospital: an international derivation and validation study. *Thorax* 2003; 58: 377–382.
4 Miyashita N, Matsushima T, Oka M, *et al.* The JRS guidelines for the management of community-acquired pneumonia in adults: an update and new recommendations. *Intern Med* 2006; 45: 419–428.
5 Charles PGP, Wolfe R, Whitby M, *et al.* SMART-COP: a tool for predicting the need for intensive respiratory or vasopressor support in community-acquired pneumonia. *Clin Infect Dis* 2008; 47: 375–384.
6 Shi SJ, Li H, Liu M, *et al.* Mortality prediction to hospitalized patients with influenza pneumonia: PO(2) /FiO(2) combined lymphocyte count is the answer. *Clin Respir J* 2017; 11: 352–360.
7 Guo L, Wei D, Zhang X, *et al.* clinical features predicting mortality risk in patients with viral pneumonia: the MuLBSTA score. *Front Microbiol* 2019; 10: 2752–2752.
8 National Health Service (NHS) England. Resources to support the safe adoption of the revised National Early Warning Score (NEWS2). www.england.nhs.uk/wp-content/uploads/2019/12/Patient_Safety_Alert_-_adoption_of_NEWS2.pdf Date last updated: 25 April 2018; date last accessed: 2 April 2020.
9 Rudd KE, Seymour CW, Aluisio AR, *et al.* Association of the Quick Sequential (Sepsis-Related) Organ Failure Assessment (qSOFA) score with excess hospital mortality in adults with suspected infection in low- and middle-income countries. *JAMA* 2018; 319: 2202–2211.

10 China NHCotpsRo. Chinese management guideline for COVID-19 (Version 6.0). www.chinadaily.com.cn/pdf/2020/1.Clinical.Protocols.for.the.Diagnosis.and.Treatment.of.COVID-19.V7.pdf Date last updated: 24 March 2020; date last accessed: 12 August 2020
11 Huang C, Wang Y, Li X, *et al.* Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* 2020; 395: 497–506.
12 Wu C, Chen X, Cai Y, *et al.* Risk factors associated with acute respiratory distress syndrome and death in patients with Coronavirus disease 2019 pneumonia in Wuhan, China. *JAMA Intern Med* 2020; 180: 1–11.
13 Xu Z, Shi L, Wang Y, *et al.* Pathological findings of COVID-19 associated with acute respiratory distress syndrome. *Lancet Respir Med* 2020; 8: 420–422.
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: 1054–1062.