

ORIGINAL ARTICLE



A risk score based on baseline risk factors for predicting mortality in COVID-19 patients

Ze Chen^{a,b,*}, Jing Chen^{b,c,*}, Jianghua Zhou^{a,b,*}, Fang Lei^{a,b,*}, Feng Zhou^{b,d}, Juan-Juan Qin^{a,b}, Xiao-Jing Zhang^{a,b}, Lihua Zhu^{a,b}, Ye-Mao Liu^{a,b}, Haitao Wang^e, Ming-Ming Chen^{a,b}, Yan-Ci Zhao^{a,b}, Jing Xie^a, Lijun Shen^{a,b}, Xiaohui Song^{a,b}, Xingyuan Zhang^{a,b}, Chengzhang Yang^{a,b}, Weifang Liu^{a,b}, Xiao Zhang^f, Deliang Guo^e, Youqin Yan^g, Mingyu Liu^h, Weiming Maoⁱ, Liming Liu^j, Ping Ye^k, Bing Xiao^l, Pengcheng Luo^m, Zixiong Zhangⁿ, Zhigang Lu^o, Junhai Wang^p, Haofeng Lu^q, Xigang Xia^r, Daihong Wang^s, Xiaofeng Liao^t, Gang Peng^u, Liang Liang^v, Jun Yang^v, Guohua Chen^w, Elena Azzolini^{x,y}, Alessio Aghemo^{x,y}, Michele Ciccarelli^{x,y}, Gianluigi Condorelli^{x,y}, Giulio G. Stefanini^{x,y}, Xiang Wei^z, Bing-Hong Zhang^{aa}, Xiaodong Huang^{ab}, Jiahong Xia^k, Yufeng Yuan^e, Zhi-Gang She^{a,b}, Jiao Guo^{ac}, Yibin Wang^{ad}, Peng Zhang^{b,d} and Hongliang Li^{a,b,d}

^aDepartment of Cardiology, Renmin Hospital, School of Basic Medical Science, Wuhan University, Wuhan, China; ^bInstitute of Model Animal, Wuhan University, Wuhan, China; ^cSchool of Mathematics and Physics, Wuhan Institute of Technology, Wuhan, China; ^dMedical Science Research Center, Zhongnan Hospital of Wuhan University, Wuhan, China; ^eDepartment of Hepatobiliary and Pancreatic Surgery, Zhongnan Hospital of Wuhan University, Wuhan, China; ^fEye Center, Renmin Hospital of Wuhan University, Wuhan, China; ^gInfections Department, Wuhan Seventh Hospital, Wuhan, China; ^hThe Ninth Hospital of Wuhan City, Wuhan, China; ⁱDepartment of General Surgery, Huanggang Central Hospital, Huanggang, China; ^jDepartment of General Surgery, Ezhou Central Hospital, Ezhou, China; ^kDepartment of Cardiology, The Central Hospital of Wuhan, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China; ^lDepartment of Stomatology, Xiantao First People's Hospital, Xiantao, China; ^mDepartment of Urology, Wuhan Third Hospital and Tongren Hospital of Wuhan University, Wuhan, China; ⁿThe Central Hospital of Enshi Tujia and Miao Autonomous Prefecture, Enshi, China; ^oDepartment of Neurology, The First People's Hospital of Jingmen affiliated to Hubei Minzu University, Jingmen, China; ^pDepartment of Orthopedics, The First People's Hospital of Jingmen affiliated to Hubei Minzu University, Jingmen, China; ^qDepartment of Hepatobiliary Surgery, The First Affiliated Hospital of Changjiang University, Jingzhou, China; ^rDepartment of Hepatobiliary Surgery, Jingzhou Central Hospital, Jingzhou, China; ^sDepartment of Hepatobiliary and Pancreatic Surgery, Xianning Central Hospital, Hubei Province, Xianning, China; ^tXiangyang Central Hospital, Affiliated Hospital of Hubei University of Arts and Science, Xiangyang, China; ^uDepartment of Hepatobiliary and Pancreatic Surgery, Suizhou Central Hospital Affiliated to Hubei Medical College, Suizhou, China; ^vDepartment of Cardiology, The First College of Clinical Medical Science, China Three Gorges University and Yichang Central People's Hospital and Institute of Cardiovascular Diseases, China Three Gorges University, Yichang, China; ^wDepartment of Neurology, Wuhan First Hospital/Wuhan Hospital of Traditional Chinese and Western Medicine, Wuhan, China; ^xHumanitas Clinical and Research Hospital IRCCS, Rozzano-Milan, Italy; ^yDepartment of Biomedical Sciences, Humanitas University, Pieve Emanuele-Milan, Italy; ^zDivision of Cardiothoracic and Vascular Surgery, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China; ^{aa}Departments of Neonatology, Renmin Hospital of Wuhan University, Wuhan, China; ^{ab}Department of Gastroenterology, Wuhan Third Hospital and Tongren Hospital of Wuhan University, Wuhan, China; ^{ac}Guangdong Metabolic Diseases Research Center of Integrated Chinese and Western Medicine and Key Laboratory of Glucolipid Metabolic Disorder, Guangdong TCM Key Laboratory for Metabolic Diseases, Ministry of Education of China and Institute of Chinese Medicine, Guangdong Pharmaceutical University, Guangzhou, China; ^{ad}Departments of Anesthesiology, Physiology and Medicine, Cardiovascular Research Laboratories, David Geffen School of Medicine, University of California, Los Angeles, CA, USA

ABSTRACT

Background: To develop a sensitive and clinically applicable risk assessment tool identifying coronavirus disease 2019 (COVID-19) patients with a high risk of mortality at hospital admission. This model would assist frontline clinicians in optimizing medical treatment with limited resources.

Methods: 6415 patients from seven hospitals in Wuhan city were assigned to the training and testing cohorts. A total of 6351 patients from another three hospitals in Wuhan, 2169 patients from outside of Wuhan, and 553 patients from Milan, Italy were assigned to three independent validation cohorts. A total of 64 candidate clinical variables at hospital admission were analyzed by random forest and least absolute shrinkage and selection operator (LASSO) analyses.

Results: Eight factors, namely, Oxygen saturation, blood Urea nitrogen, Respiratory rate, admission before the date the national Maximum number of daily new cases was reached, Age, Procalcitonin, C-reactive protein (CRP), and absolute Neutrophil counts, were identified as having significant associations with mortality in COVID-19 patients. A composite score based on these eight risk factors, termed

ARTICLE HISTORY









Received 25 November 2020

Revised 13 March 2021


Accepted 15 March 2021

KEYWORDS

COVID-19; risk score; modeling; mortality; in-hospital

CONTACT Jiao Guo  guoj@gdpu.edu.cn  Institute of Chinese Medicine, Guangdong Pharmaceutical University, 280 Wai Huan Dong Road, Guangzhou 510006, China; Yibin Wang  yibinwang@mednet.ucla.edu  Departments of Anesthesiology, Physiology and Medicine, Cardiovascular Research Laboratories, David Geffen School of Medicine, University of California, CHS 37-200J, Los Angeles, 90095 CA, USA; Peng Zhang  zhp@whu.edu.cn  Medical Science Research Center, Zhongnan Hospital of Wuhan University, 169 Donghu Road, Wuhan 430071, China; Hongliang Li  lihl@whu.edu.cn  Department of Cardiology, Renmin Hospital of Wuhan University, 99 Zhangzhidong Road, Wuhan 430060, China; Institute of Model Animal of Wuhan University, 169 Donghu Road, Wuhan 430071, China

*These authors contributed equally: Ze Chen; Jing Chen; Jianghua Zhou; Fang Lei

 Supplemental data for this article is available online at <https://doi.org/10.1080/03007995.2021.1904862>.

the OURMAPCN-score, predicted the risk of mortality among the COVID-19 patients, with a C-statistic of 0.92 (95% confidence interval [CI] 0.90–0.93). The hazard ratio for all-cause mortality between patients with OURMAPCN-score >11 compared with those with scores ≤ 11 was 18.18 (95% CI 13.93–23.71; $p < .0001$). The predictive performance, specificity, and sensitivity of the score were validated in three independent cohorts.

Conclusions: The OURMAPCN score is a risk assessment tool to determine the mortality rate in COVID-19 patients based on a limited number of baseline parameters. This tool can assist physicians in optimizing the clinical management of COVID-19 patients with limited hospital resources.

Introduction

The coronavirus disease 2019 (COVID-19) pandemic has imposed a significant burden on healthcare systems worldwide, particularly in major hotspot cities^{1,2}. While the spread of infections appears to be abating in some countries, a growing COVID-19 crisis is still emerging in the developing world, where the medical resources and infrastructure are far less advanced and can be easily overwhelmed³. There is an urgent need to develop readily applicable triage criteria to identify COVID-19 patients with a poor prognosis at admission to enable the allocation of the limited medical resources to those who most need them^{2,4}. Currently, the disease severity and the risk of death in patients with COVID-19 are often graded with the use of either a single clinical variable^{5–7}, such as lymphopenia or hypoxemia, or by doctors' empirical evaluations⁸. Identifying COVID-19 patients at high risk of mortality based on a sensitive and quantitative risk assessment tool would be valuable for optimizing care and reducing mortality⁸. Recently, ISARIC4C investigators have developed a 4C mortality score stratify patients admitted to hospital with COVID-19 into the different risk of mortality based on the European population⁹. However, COVID-19 is still widely spreading in Asian countries.

Here, we report the development and validation of a risk assessment tool using readily accessible clinical variables at baseline based on an observational outcome study from a large cohort of COVID-19 in-hospital patients in Hubei Province, China. An integrated risk score (OURMAPCN score) was derived from eight risk factors, namely, **O**xygen saturation, blood **U**rea nitrogen (BUN), **R**espiratory rate, admission before the date the first national **M**aximum number of daily new cases was reached, **A**ge 60 and above, **P**rocalcitonin, **C**-reactive protein (CRP) and absolute **N**eutrophil counts. We showed that a higher OURMAPCN score predicted a higher mortality rate in COVID-19 patients with adequate specificity and sensitivity across the training, testing, and three different validation cohorts, including two from China and one from Italy. Therefore, the OURMAPCN score has satisfactory performance with regard to predicting the risk of mortality across different clinical cohorts. A website dedicated to the OURMAPCN score (<http://compute.covid-ourmap.cn:8888/comput-OURMAPCN>) will support frontline physicians performing risk stratification among COVID-19 patients. We propose that the application of the OURMAPCN score could assist physicians in accurately assessing the mortality risk in COVID-19 patients at admission to optimize management options, particularly in areas with limited resources in Asian countries.

Methods

Inpatient cohorts and study procedure

We performed a retrospective observational study with a total of 15,488 inpatients with confirmed COVID-19 in Hubei Province, China and Milan, Italy. As COVID-19 cases in China were concentrated in January to March in 2020, and the number of daily new cases was very limited after March, we included COVID-19 cases mainly admitted to hospitals from January to March, 2020 in Hubei cohort.

For the training and testing cohorts, a total of 6415 patients with confirmed COVID-19 who were admitted between 1 January 2020, and 20 March 2020, were consecutively included in the study from seven COVID-19 designated hospitals in Wuhan city. The final date of follow-up for outcome determination was 8 April 2020. A total of 611 patients remaining in the hospital at the end of follow-up were treated as censor in Cox models. No exclusion criterion was applied in this study. A total of 462 patients died out of the 6415 patients during hospitalization. Among these patients, 70% (4492) were randomly assigned to the training cohort, and the remaining 30% (1923) patients were assigned to the testing cohort (Figure 1 and Supplemental Table 1).

There were 6351 patients with COVID-19 who were admitted between 1 January 2020 and 17 April 2020, in another three designated hospitals in Wuhan city; these patients were consecutively enrolled in the study and were assigned to validation cohort 1 (Figure 1 and Supplemental Table 1). The final date of the follow-up was 26 April 2020. A total of 257 patients remaining in the hospital at the end of follow-up were treated as censor in Cox models. No exclusion criterion was applied in this study. A total of 587 patients died out of the 6351 patients during hospitalization.

A total of 2169 patients with COVID-19 from eleven hospitals in Hubei Province outside of Wuhan city were designated as validation cohort 2 (Figure 1 and Supplemental Table 1). The inclusion of this subset of patients started from 1 January 2020, and ended on 9 March 2020; patients were enrolled consecutively. The last follow-up date was 27 March 2020. 18 patients remaining in the hospital at the end of follow-up were treated as censor in Cox models. No exclusion criterion was applied to the study. A total of 132 patients died during hospitalization out of the 2169 patients.

The Italian cohort (validation cohort 3) initially enrolled 553 patients with confirmed COVID-19 admitted between 12 February 2020, and 12 April 2020, in Humanitas Research Hospital in Milan, Italy (Figure 1). The patients were enrolled consecutively. The last follow-up date was 30 April. At the

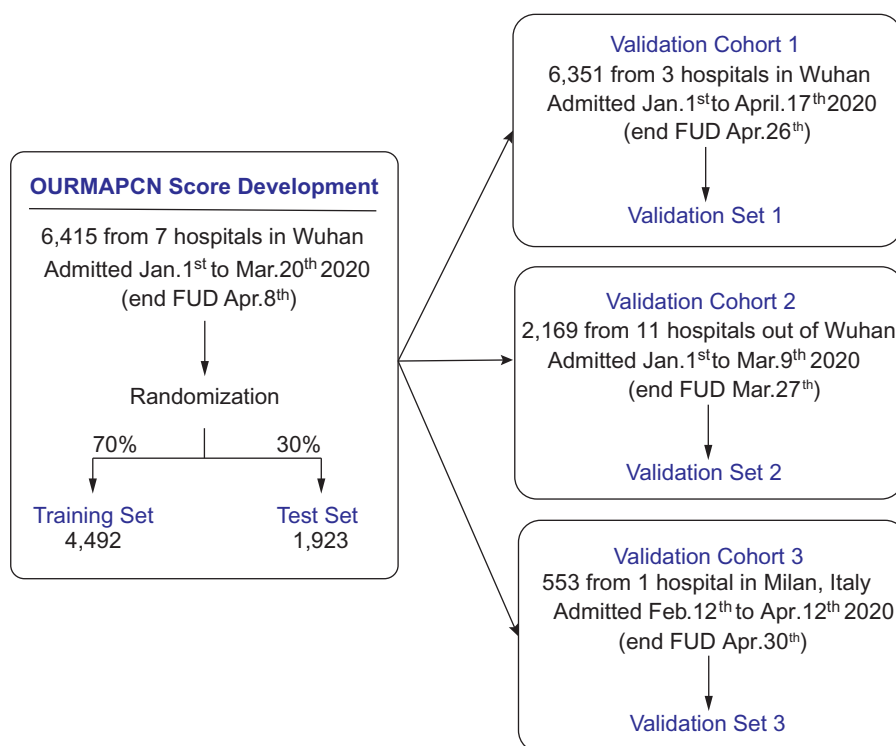


Figure 1. Flowchart for patient selection and distribution of the training, the test and the validation cohorts. FUD indicates the last of follow-up date.

end of the study, 24 patients remained in the hospital and were treated as censor in Cox models. A total of 124 patients died during hospitalization out of the 553 patients.

COVID-19 was diagnosed by clinical manifestations, chest CT, or real-time RT-PCR according to the World Health Organization (WHO) interim guidance and/or the New Coronavirus Pneumonia Prevention and Control Program (5th edition) published by the National Health Commission of China^{10,11}. The demographics, clinical characteristics, medical history, laboratory tests, radiological reports, therapeutic intervention, and outcome data were obtained from patients' electronic medical records. All variables used for establishing the predictive models were based on the measurements at admission and not at another time point during hospitalization in the training and test cohorts. Measurements at admission were also used in all validation cohorts. Patients transferred between hospitals were not included in our study. The ethics committee of each hospital approved the study protocols. The need to obtain informed consent from the patients was waived by each ethics committee.

Preparation of candidate variables

We followed the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD) statement for reporting multivariable prediction model development and validation. We identified candidate variables for potential inclusion in our models from a review of the literature and identified clinical variables available at hospital admission in COVID-19 patients, including patients' demographics, clinical presentation, medical history, laboratory tests and radiological examinations, outcomes, admission date and hospital sites. All major software and code

used to analyze these datasets are referenced in [Supplemental Table 2](#).

Initially, a total of 162 variables on admission were collected from 6415 patients hospitalized with COVID-19 in Wuhan. The variables with a large proportion of missing (>50%) or with a correlation of 0.9 or higher were not included for variable selection process. Categorical variables with small categories (less than 5%) were either combined categories or removed. Continuous variables from different institutions were scaled by site and gender and categorical variables were categorized with/without their intrinsic ordering. Therefore, the variables for the risk prediction model were selected from a total of 64 candidate clinical parameters through a two-round process ([Figure 2](#)). Before the variable selection process, this Wuhan cohort was randomly split at a ratio of 70%/30% into the training (4492) and testing (1923) cohorts, respectively, and multiple imputations were used to account for missing data in these two separate data sets¹².

Variable selection using random forest

In the first-round of variable selection, a random forest model was applied in the training dataset. This machine learning method works on both categorical and numerical variables on any scale, obviating the need for the conversion of features or the normalization^{13,14}. The random forest model is also robust with regard to controlling potential overfitting and is adequately parsimonious in a large dataset¹⁵⁻¹⁷. To avoid overfitting, a pre-pruning algorithm was set for each node to stop the process if all variables belonged to the same class or if all the variable values were the same^{18,19}. Additionally, the number of tree was set at

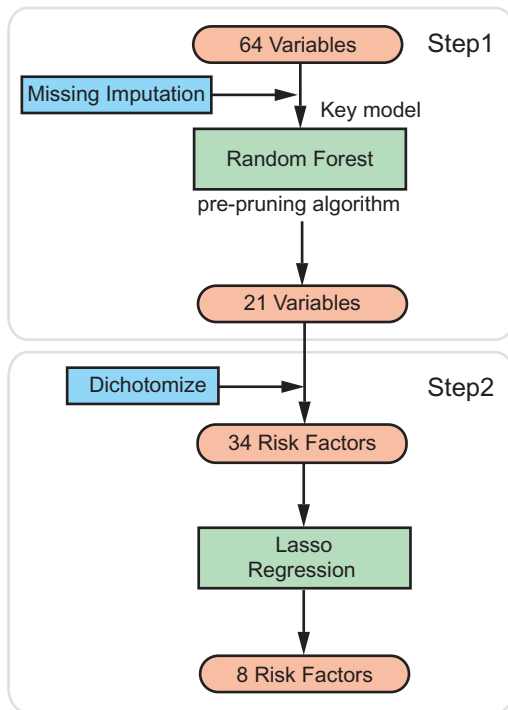


Figure 2. Flowchart for variable selection. The random forest algorithm was applied to identify variables that may influence prediction among 64 clinical variables collected on admission. 21 variables were derived from the random forest analysis. For better clinical applicability, the 21 variables were dichotomized into another 34 risk factors for a Least Absolute Shrinkage and Selection Operator (LASSO) regression analysis. A total of eight risk factors at admission were extracted.

500, and the number of variables randomly sampled at each split was optimized with respect to Out-of-Bag error estimates ($mtry = 6$)¹⁴. Then, variables were dropped one by one with the least relevant variable until the mean decreased accuracy larger than 0.001²⁰. The selected variables were ranked by variable importance measures (VIMs) to identify those that would most significantly influence the prediction^{21,22}.

Variables dichotomization and candidate risk factors selection using LASSO model

To improve the clinical utility, all variables selected in the random forest model were dichotomized. The abnormal cut-off for dichotomization was applied according to the reference range in each hospital (Supplemental Table 3). The date of the first peak in the relevant country's daily number of new cases was obtained from the WHO and worldometers.info (Supplemental Table 4). We referred to the dichotomized variables as risk factors.

To further select candidate risk factors in the second step, a Least Absolute Shrinkage and Selection Operator (LASSO) regression was applied to minimize the potential collinearity of variables and overfitting (Figure 2). The glmnet package in R was applied to efficiently search for sparse solutions and to optimize the solutions using a 10-fold cross-validation schema over the training dataset²³ and in the study, λ ($=0.013$) is chosen as the largest lambda at which the mean squared error (MSE) is within one standard error of the

minimal MSE by 10 times cross validation. A final LASSO model was selected from the most regularized model which had a within one standard error of the minimum (Figure 2).

The OURMAPCN score derivation

In the score building stage, a multivariate Cox hazard regression analysis was employed for the final list of risk factors from the second round of selection. The assignment of points to risk factors was based on a rounded number that was approximately four times of the coefficient in the Cox model with the maximal value set at five.

Model validation and calibration

The bootstrapped C-statistic was applied to validate the performance of the Cox hazard model as well as calibration plot. Separate C-statistics were analyzed for OURMAPCN risk score in the training, test, and validation groups. Additionally, datasets from all groups were split equally into several groups. For a dataset with fewer events than the controls, we created balanced splits of the data in all groups by maintaining the outcome structure for the training and testing datasets with a 70%/30% ratio in total of cases.

Missing data imputation and cross validation

Because multivariate analysis requires a complete set of variables for each patient, missing data from one or more non-invasive tests were imputed in the study. Multiple imputations were implemented to handle missing data by the random forest package "missForest" algorithm, which is a highly accurate method for lab data^{12,24,25}. Mixed types of data (continuous and categorical) were imputed in a parallel process. All variables entering the random forest model were included in the imputation models. Missing patterns were compared across age and sex to ensure that missing data were not correlated with basic personal characteristics. The variables with a large proportion of missing ($>50\%$) were excluded from the random forest model variable selection process. The MissForest algorithm was applied to the missing parameters to estimate the imputation error as follows: 1.30 for the normalized root mean squared error and 5.00% for false classification²⁵. A bootstrapped cross validation process was applied to the training data by artificially and randomly introducing 10% more missing data, and after 10 iterations, MissForest yielded 4.21% (IQR 3.60%–4.90%) differences in the continuous parameters and 0.17% (IQR 0.15–0.18%) differences in the categorical parameters when the two datasets with or without artificially introduced missing parameters were compared. In brief, the levels of missing data for the selected parameters in final score building were as follows: 0% for age, 0% for admission date, 5.48% for respiratory rate, 15.05% for absolute neutrophil count, 15.76% for BUN, 18.14% for oxygen saturation, 31.97% for CRP, and 30.00% for procalcitonin.

Sensitivity analyses

A sensitivity analysis was performed for all the variable/factor selection processes on the complete cases without imputation. We also conducted a sensitivity analysis using univariate general linear model (GLM) followed by pair-matching comparison and random forest model in step 1, and GLM stepwise backward model in step 2, to select candidate variables. A multivariate Cox proportional hazard model was used to calculate the risk score in this sensitivity analysis.

Results

Patient characteristics in the training and testing cohorts

To determine the risk factors significantly associated with in-hospital mortality, a total of 6415 confirmed COVID-19 patients from seven hospitals in the city of Wuhan, Hubei Province, were included and assigned to the training and test cohorts at a 70%/30% ratio (see details in Section "Methods" and Figure 1). Their baseline characteristics are listed in Table 1. The median age of the participants was 59 years (interquartile range [IQR], 46–68), the median respiratory rate was 20 (IQR, 18–21), the median oxygen saturation (SpO₂) was 98% (IQR, 96–98) and the median follow-up duration was 15 days (IQR, 10–23). A total of 1208 (25.5%) patients had elevated procalcitonin levels, and 2329 (49.5%) patients had increased CRP levels. A total of 535 (9.6%) patients showed a higher level of BUN, and 819 (14.5%) patients had an elevated neutrophil count. A total of 3836 (59.8%) patients were admitted to the hospital before 12 February (the date of the peak number of daily new cases in China). A total of 462 patients died in the study period. The differences in baseline parameters between the patients who survived and the patients who died in the training and testing cohort are shown in Table 1.

OURMAPCN score and Cox analysis

After the random forest model selection, 21 clinical variables were selected among 64 clinical parameters recorded at admission based on the patient's data from training cohort (Supplemental Table 5). To increase clinical applicability, the 21 variables were dichotomized into another 34 risk factors and further evaluated by a LASSO algorithm, leading to the extraction of eight risk factors at admission that were closely associated with mortality in COVID-19 patients (Supplemental Table 6). The final eight risk factors, namely, Oxygen saturation < 90%, blood Urea nitrogen > upper limit of normal (ULN), Respiratory rate > 30, admission before the date of the first national Maximum number of daily new cases was reached, Age ≥ 60, Procalcitonin > ULN, C-reactive protein (CRP) > ULN and absolute Neutrophil counts > ULN, were incorporated into the risk score model for the prediction of mortality in COVID-19 patients. Sensitivity analysis with the complete cases without imputation obtained seven of the eight variables except procalcitonin > ULN.

To calculate the risk score, we assigned each of the eight prognostic factors a numeric point value that was

proportional to its specific regression coefficient in a multivariate Cox proportional hazard model. Calibration curves for predicting probabilities of survival at 14 days, 21 days, and 28 days were computed from the Cox models on which the risk score based²⁶. We observed a high degree of similarity between the observed and the estimated rate at 14 days, 21 days, and 28 days (Supplemental Figure 1). Patients who remained in the hospital at the end of follow-up time were treated as right censor. Scores were calculated by summing the points for each of the eight factors; the points for each factor were obtained by multiplying the coefficient in the Cox model by four and then rounding, with a maximum value set at five (Table 2).

We also conducted a sensitivity analysis using univariate GLM followed by pair-matching comparison and random forest model in step 1, and GLM stepwise backward model in step 2, to select candidate variables. A multivariate Cox proportional hazard model was used to calculate the risk score. As shown in Supplemental Table 7, the variables in the final risk score in sensitivity analysis were the same as the OURMAPCN score.

Prediction of mortality in the training and test cohorts

The OURMAPCN score had a high prognostic significance in the training cohort, with an AUROC of 0.92 (95% CI, 0.90–0.93) (Table 3 and Supplemental Figure 2). A density plot was generated to determine a cutoff value for the OURMAPCN score to discriminate the patients with low and high risks of mortality: the low-risk group had an OURMAPCN score from 0 to 11 points, and a high-risk group had an OURMAPCN score from 12 to 23 points (Supplemental Figure 3). The Kaplan–Meier analysis showed that the patients with scores above 11 points had a significantly decreased survival rate in the training cohort, with an HR of 18.18 (95% CI, 13.93–23.71; $p < .0001$) (Supplemental Figure 4(A)). In the test cohort, the OURMAPCN score also had high C statistic indexes, with AUROCs of 0.90 (95% CI, 0.87–0.92) (Table 3 and Supplemental Figure 2). The HR for mortality was 12.36 (95% CI, 8.79–17.38; $p < .0001$) in patients with scores of more than 11 points in the test cohort (Supplemental Figure 4(A)). A webpage dedicated to OURMAPCN score calculation (<http://compute.covid-ourmap.cn:8888/comput-OURMAPCN>) has been developed to support doctors at the point-of-care.

OURMAPCN score performance in the external validation cohorts

The patients' characteristics in validation cohort 1, which consisted of 6351 patients from three independent hospitals in Wuhan city, are detailed in Supplemental Table 8. The OURMAPCN score had an AUROC of 0.89 (95%CI, 0.88–0.91) in the validation cohort 1 (Supplemental Figure 2) (Table 4). Based on the OURMAPCN score, patients with scores greater than 11 points had a higher risk of mortality than those with lower scores, with an HR of 14.35 (95% CI, 11.00–18.73; $p < .0001$) (Supplemental Figure 4(B)).

Table 1. Patients' characteristics of the training and test cohorts.

Variables	All (6415)	Survived (5342)	Died (462)	p Value ^d
Clinical characteristics on admission				
Median age (IQR)—yr	59 (46–68)	57 (45–66)	71 (63–80)	<.001
Male sex—no./total no. (%)	3028/6415 (47.2)	2442/5342 (45.7)	291/462 (63.0)	<.001
Median heart rate (IQR)—bpm ^a	84 (78–96)	84 (78–96)	89 (79–104)	<.001
Median respiratory rate (IQR)—bpm ^a	20 (18–21)	20 (18–21)	21 (20–26)	<.001
Median SBP (IQR)—mmHg	129 (120–140)	128 (120–140)	128 (118–142)	.91
Median DBP (IQR)—mmHg	79 (72–86)	79 (72–86)	76 (69–85)	<.001
Fever—no./total no. (%)	4466/6139 (72.7)	3728/5160 (72.2)	343/421 (81.5)	<.001
Median follow-up time (IQR)—days	15 (10–23)	16 (10–23)	8 (4–15)	<.001
Comorbidities on admission				
Chronic obstructive pulmonary disease—no./total no. (%)	75/6414 (1.2)	60/5342 (1.1)	9/461 (2.0)	.12
Diabetes—no./total no. (%)	806/6414 (12.6)	613/5342 (11.5)	97/461 (21.0)	<.001
Coronary heart disease—no./total no. (%)	460/6414 (7.2)	328/5342 (6.1)	85/461 (18.4)	<.001
Cerebrovascular diseases—no./total no. (%)	204/6414 (3.2)	138/5342 (2.6)	40/461 (8.7)	<.001
Hypertension—no./total no. (%)	1935/6414 (30.2)	1496/5342 (28.0)	234/461 (50.8)	<.001
Laboratory examination on admission				
Neutrophil count > 6.3 × 10 ⁹ /L—no./total no. (%)	819/5630 (14.5)	493/4656 (10.6)	227/423 (53.7)	<.001
Lymphocyte count < 1.1 × 10 ⁹ /L—no./total no. (%)	2193/5631 (38.9)	1579/4656 (33.9)	351/423 (83.0)	<.001
Platelet count < 125 × 10 ⁹ /L—no./total no. (%)	539/5637 (9.6)	347/4662 (7.4)	131/423 (31.0)	<.001
C-reactive protein > ULN—no./total no. (%) ^b	2329/4706 (49.5)	1773/3978 (44.6)	334/341 (97.9)	<.001
Procalcitonin > ULN—no./total no. (%) ^b	1208/4740 (25.5)	783/3875 (20.2)	268/378 (70.9)	<.001
BUN > ULN—no./total no. (%) ^b	535/5593 (9.6)	280/4630 (6.0)	186/421 (44.2)	<.001
Total cholesterol > 5.17 mmol/L—no./total no. (%)	609/4899 (12.5)	537/4022 (13.4)	13/374 (3.5)	<.001
D-dimer > ULN—no./total no. (%) ^b	2106/4760 (44.2)	1516/3897 (38.9)	316/380 (83.2)	<.001
Low density lipoprotein > 3.37 mmol/L—no./total no. (%)	608/4899 (12.4)	509/4022 (12.7)	23/374 (6.1)	<.001
Median SpO ₂ (IQR)—%	98 (96–98)	98 (96–98)	91 (81–96)	<.001
Other risk factors				
Teaching hospital—no./total no. (%)	3359/6415 (52.4)	2532/5342 (47.4)	309/462 (66.9)	<.001
Admission before Feb 12th—no./total no. (%) ^c	3836/6415 (59.8)	2004/5342 (37.5)	356/462 (77.1)	<.001

Abbreviations. SBP, systolic blood pressure; DBP, diastolic blood pressure; BUN, blood urea nitrogen; SpO₂, oxygen saturation; ULN, upper limit of normal.

^aThe bpm in heart rate and respiratory rate denotes beat per minute and breath per minute, respectively.

^bULN indicates the upper limit of the normal range of each biochemical test. The reference ranges of the test in each hospital were provided in [Supplemental Table 3](#).

^cThe patients who admitted to hospital before 12 February 2020 when the daily newly diagnosed cases start to decline in China.

^dp Values for the comparison of survivors with patients who died and were calculated by the Mann-Whitney U test for non-normally distributed continuous variables and by the chi-square test or Fisher's exact test for categorical variables.

Table 2. Cox model for the scores of selected risk factors.

Risk factor	Incidence of death % ^c	Cox model		
		HR (95% CI) ^d	Coefficient	Points ^e
C-reactive protein > ULN ^a	11.02	6.50(3.68–11.48)	1.87(1.30–2.44)	5
SpO ₂ < 90%	54.98	3.17 (2.37–4.23)	1.15(0.86–1.44)	4
Admission before Feb 12th ^b	13.38	1.85 (1.41–2.43)	0.61(0.34–0.89)	3
Age ≥ 60	20.41	1.99 (1.42–2.78)	0.69(0.35–1.02)	3
BUN > ULN ^a	35.62	2.29 (1.79–2.94)	0.83(0.58–1.08)	2
Respiratory rate ≥ 30	43.00	2.57(1.64–4.04)	0.94(0.49–1.40)	2
Procalcitonin level > ULN ^a	14.62	2.10 (1.61–2.72)	0.74(0.48–1.00)	2
Neutrophil count > 6.3 × 10 ⁹ /L	29.47	2.11 (1.65–2.70)	0.75(0.50–0.99)	2

Abbreviations. BUN, blood urea nitrogen; SpO₂, oxygen saturation; ULN, upper limit of normal; HR, hazard ratios.

^aThe ULN denotes the upper limit of the normal range of each biochemical test. The reference ranges of tests in each hospital were provided in [Supplemental Table 3](#).

^bThe proportion of patients who admitted to hospital before 12 February 2020 when the daily newly diagnosed cases start to decline in China.

^cThe incidences of death in each risk factor were calculated by the number of deaths to the total number of patients with each increased risk.

^dHazard ratios with corresponding 95% confidence intervals (CIs) were calculated by multivariate Cox regression analysis and were for the comparison of survivors with patients who died.

^eThe assignment of points to risk factors was based on a multivariable Cox regression coefficient.

Another cohort of 2169 patients with COVID-19 from eleven hospitals outside of Wuhan city was designated validation cohort 2. The baseline characteristics of the patients in the validation cohort 2 are described in [Supplemental Table 9](#). The OURMAPCN score had an AUROC of 0.90 (95%CI, 0.88–0.93) ([Supplemental Figure 2](#)) ([Table 4](#)) in the validation cohort 2. Again, patients with an OURMAPCN score greater than 11 points had a higher risk of mortality than those with lower scores, with an HR of 32.14 (95% CI, 7.95–129.9; $p < .0001$) ([Supplemental Figure 4\(C\)](#)).

The OURMAPCN score (AUROC 0.89 [95%CI, 0.88–0.91] in validation cohort 1 and AUROC 0.90 [95%CI, 0.88–0.93] in validation cohort 2) significantly outperformed the existing scoring systems, including the MuLBSTA (AUROC 0.58 [95%CI, 0.56–0.60] in validation cohort 1 and AUROC 0.79 [95%CI, 0.75–0.83] in validation cohort 2) and CURB-65 (AUROC 0.81 [95% CI, 0.79–0.83] for validation cohort 1 and AUROC 0.80 [95% CI, 0.76–0.84] in validation cohort 2) scores for the prediction of mortality in COVID-19 patients ([Table 5](#)).

Table 3. Performance of OURMAPCN score in training and test dataset.

	OURMAPCN score in training dataset	OURMAPCN score in test dataset
N	4063	1741
AUROC (95% CI)	0.92 (0.90–0.93)	0.90 (0.87–0.92)
Cutoff value	≤11	≤11
Total accuracy, % (95% CI)	77.48 (76.00–89.54)	80.59 (78.58–82.83)
Sensitivity, % (95% CI)	92.36 (77.39–95.54)	87.16(81.08–91.89)
Specificity, % (95% CI)	76.18 (74.58–90.34)	79.97 (77.84–82.49)
PPV, % (95% CI)	24.66 (23.31–41.03)	28.81 (26.36–31.77)
NPV, % (95% CI)	99.16 (97.96–99.51)	98.54 (97.85–99.08)
PLR, (95% CI)	3.86 (3.61–4.12)	4.47 (3.97–5.03)
NLR, (95% CI)	0.10 (0.07–0.14)	0.16 (0.10–0.23)

Abbreviations. AUROC, area under the receiver operating characteristics; PPV, positive predictive value; NPV, negative predictive value; PLR, positive likelihood ratio; NLR, negative likelihood ratio; 95% CI, 95% confidence interval.

Table 4. OURMAPCN score performance in validation cohorts in Wuhan, other cities in Hubei, China, and Milan, Italy.

	Wuhan validation cohort 1	Other cities in Hubei validation cohort 2	Milan, Italy validation cohort 3
AUROC (95% CI)	0.89 (0.88–0.91) ^a	0.90 (0.88–0.93) ^a	0.81 (0.76–0.86) ^a
Total accuracy, % (95% CI)	89.07 (72.74–89.92)	83.45 (80.80–90.10)	77.10 (62.81–84.35)
Sensitivity, % (95% CI)	72.57 (68.82–91.31)	86.36 (76.52–92.42)	67.82 (51.72–88.51)
Specificity, % (95% CI)	90.92 (70.93–91.81)	83.06 (80.39–90.74)	79.38 (57.63–90.11)
PPV, % (95% CI)	45.70 (24.68–48.50)	25.65 (22.29–36.31)	44.70 (32.58–60.00)
NPV, % (95% CI)	96.89 (96.49–98.73)	98.92 (98.32–99.41)	91.12 (88.14–95.58)
PLR, (95% CI)	7.99 (7.25–8.82)	5.14 (4.56–5.79)	3.29 (2.57–4.27)
NLR, (95% CI)	0.30 (0.27–0.35)	0.16 (0.10–0.24)	0.40 (0.28–0.54)

Abbreviations. AUROC, area under the receiver operating characteristics; PPV, positive predictive value; NPV, negative predictive value; PLR, positive likelihood ratio; NLR, negative likelihood ratio; 95% CI, 95% confidence interval.

^aThe C statistic metrics for OURMAPCN score were reported in validation cohorts in Wuhan, other cities in Hubei, China and Milan, Italy.

Table 5. Comparison of OURMAPCN score with the MuLBSTA, CURB-65, and COVID-GRAM scores for the identification of high risk of mortality.

Scores	Validation 1: Wuhan		Validation 2: Other cities in Hubei	
	N	AUROC (95% CI)	N	AUROC (95% CI)
OURMAPCN score	6094	0.89 (0.88–0.91)	2151	0.90 (0.88–0.93)
MuLBSTA score ^a	6094	0.58 (0.56–0.60)	2151	0.79 (0.75–0.83)
CURB-65 score ^b	6094	0.81 (0.79–0.83)	2151	0.80 (0.76–0.84)
COVID-GRAM ^c	6094	0.87 (0.86–0.89)	2151	0.85 (0.82–0.89)

Abbreviations. AUROC, area under the receiver operating characteristics;

^aMuLBSTA, a score used in patients with viral pneumonia, to predict clinical characteristics that affect mortality. It comprises factors: multilobe infiltrate (5), absolute lymphocyte counts $\leq 0.8 \times 10^9/L$ (4), bacterial infection (4), acute smoker (3), quit smoking (2), hypertension (2), and age ≥ 60 years (2).

^bCURB-65, a score used in the emergency department setting to risk stratify a patient's community-acquired pneumonia. It comprises factors: confusion (1), BUN above 7 mmol/L (1), respiratory rate ≥ 30 (1), systolic BP < 90 mmHg or diastolic BP ≤ 60 mmHg (1), and Age ≥ 65 (1).

^cCOVID-GRAM, a score used in patients with COVID-19, to predict the risk of developing critical illness. It comprises factors: X-ray abnormality, age, hemoptysis, dyspnea, unconsciousness, no. of comorbidities, cancer history, neutrophil to lymphocyte ratio, lactate dehydrogenase, and direct bilirubin.

When compared with a recently published COVID-19 risk score COVID-GRAM²⁷, the OURMAPCN score has a similar performance in the Wuhan population and a superior performance in patients outside of Wuhan city (Table 5).

Importantly, in addition to the evaluation of the scoring system in the training and the test cohorts, the predictive performance of the OURMAPCN score was evaluated in the subgroup of patients with different age ranges, sexes, and medical histories of chronic obstructive pulmonary disease (COPD), coronary artery disease (CAD), type 2 diabetes mellitus (T2DM) and hypertension (Table 6). The performance of the OURMAPCN score was robust and stable across different subsets of patients, with AUROCs ranging from 0.79 to 0.92.

OURMAPCN score performance in the Italian cohort

To further explore the performance of the OURMAPCN score in patients outside of China, we calculated the C-statistic in a cohort of COVID-19 patients from Milan, Italy. The baseline characteristics of this cohort are described in Supplemental

Table 10. In this Italian cohort, serum urea levels were not commonly measured in patients with COVID-19. We used serum creatinine as a substitute risk factor for BUN. The modified OURMAPCN score had an AUROC of 0.81 (95%CI, 0.76–0.86), and the other C-statistic metrics are listed in Supplemental Figure 2 and Table 4. Therefore, the OURMAPCN score yielded a satisfactory accuracy in the Italian cohort.

Discussion

We have developed the OURMAPCN score using eight baseline risk factors at admission, namely, age, respiratory rate, oxygen saturation, absolute neutrophil counts, CRP, BUN, procalcitonin, and admission before the date of the first maximum number of daily new cases in the country. The OURMAPCN score is a robust and reliable tool for predicting mortality in patients with COVID-19. These eight factors are based on common clinical parameters routinely obtained in the emergency room or at hospital admission. Therefore, the

Table 6. Performance of OURMAPCN score for subgroups in validation cohorts from Wuhan and other cities in Hubei.

Subgroups ^a	Mortality rate %	AUROC (95% CI)	Sensitivity% (95% CI)	Specificity% (95% CI)	PPV% (95% CI)	NPV% (95% CI)
Age						
<65	4.28	0.90 (0.88–0.92)	83.26 (77.68–87.98)	85.32 (84.33–88.38)	20.30 (18.87–23.84)	99.13 (98.84–99.38)
≥65	17.36	0.84 (0.82–0.86)	72.63 (68.72–76.54)	83.10 (81.59–84.66)	47.42 (44.83–50.13)	93.52 (92.69–94.42)
Gender						
Female	5.60	0.90 (0.89–0.92)	88.09 (83.40–94.04)	78.03 (71.78–79.60)	18.96 (16.01–20.39)	99.11 (98.76–99.52)
Male	11.97	0.89 (0.88–0.91)	87.60 (69.21–91.53)	73.41 (71.22–91.35)	31.48 (29.50–53.26)	97.73 (95.60–98.45)
COPD						
No	7.83	0.92 (0.90–0.93)	87.98 (84.14–91.56)	82.14 (78.18–83.44)	29.40 (25.87–31.21)	98.77 (98.40–99.11)
Yes	35.16	0.87 (0.79–0.94)	84.38 (56.25–96.88)	77.97 (59.32–96.61)	67.44 (54.54–91.30)	89.36 (79.71–97.78)
CAD						
No	7.74	0.91 (0.90–0.92)	89.39 (72.84–92.02)	76.91 (75.77–93.13)	24.55 (23.51–47.56)	98.85 (97.60–99.13)
Yes	20.41	0.79 (0.74–0.83)	80.00 (50.77–88.46)	62.92 (55.23–88.17)	36.03 (32.04–55.78)	92.09 (87.35–95.23)
Diabetes						
No	7.64	0.91 (0.90–0.92)	89.31 (86.49–91.74)	77.67 (76.66–78.66)	24.87 (23.87–25.87)	98.88 (98.59–99.13)
Yes	14.67	0.85 (0.81–0.88)	69.35 (62.90–84.95)	85.49 (67.74–87.71)	45.08 (30.49–49.62)	94.19 (93.02–96.31)
Hypertension						
No	5.97	0.91 (0.90–0.93)	87.31 (83.38–91.84)	81.81 (77.59–83.06)	23.12 (20.08–24.77)	99.02 (98.72–99.35)
Yes	14.36	0.86 (0.84–0.88)	71.91 (67.27–76.29)	86.17 (84.70–87.55)	46.52 (43.53–49.57)	94.82 (93.98–95.59)

Abbreviations. AUROC, area under the receiver operating characteristics; PPV, positive predictive value; NPV, negative predictive value; COPD, chronic obstructive pulmonary disease; CAD, coronary artery disease.

^aThe C statistic for OURMAPCN score is reported in age, gender, COPD, CAD, diabetes, and hypertension subgroup population among all study population.

OURMAPCN score can be readily determined and implemented as a very practical tool for patient risk stratification, not only by intensive care specialists but also by physicians with diverse backgrounds and specialties. This robust yet easy-to-implement clinical risk score will be valuable in the dedicated COVID-19 clinics, particularly when medical resources are limited in Asian population during the COVID-19 pandemic.

At the functional level, the OURMAPCN score integrates multidimensional information. One domain represented by the model is the degree of pulmonary impairment (respiratory rate and oxygen saturation)²⁸; one may reflect the status of the patient's systemic response to infection and the ensuing inflammation (absolute neutrophil counts²⁹, CRP³⁰ and procalcitonin³¹), and another domain (BUN^{32,33}) may indicate the level of tissue damage and breakdown and renal function. Age is recognized as a major risk factor that is positively associated with the risk of mortality due to COVID-19³⁴. Finally, this cohort comprises a large proportion of patients at the start of the pandemic when the experience of treating patients with SARS-CoV-2 infection and medical resources were limited (before the date of the first maximum number of daily new cases). It is intuitive that limited medical resources and lack of experiences can lead to a lack of life-saving intervention for critically ill COVID-19 patients. Fortunately, the majority of countries have reached the first peak and with accumulated knowledge in dealing with COVID-19 patients (Supplemental Table 4). Thus, these parameters, which were selected through an unbiased approach in a large COVID-19 cohort may offer some valuable insights into the potential pathogenic mechanisms underlying the poor outcome in the subset of COVID-19 patients.

There are several other baseline clinical variables that have been reported to be associated with COVID-19 severity and mortality, including BMI, COVID-19 symptoms (e.g. fever, short of breath), D-Dimer, and troponin^{35–37}. Indeed, we have included BMI, fever, and D-Dimer in our original set of variables. However, these variables were ruled out in the first-round of variable selection using a random forest model.

The potential explanation for BMI has not been selected by the model might be associated with the lower BMI and narrow BMI distribution in the Chinese population. The narrow distribution of a predictor limited its ability to capture its impact on an outcome. Because short of breath had a high correlation with respiratory rate, to avoid overfitting, we only applied respiratory rate in our variable selection process. As for troponin, this test was chosen only when the patients were suspected of having a severe cardiac injury or cardiac ischemia. The proportion of available values only counted up 5% for cardiac troponin I and 10% for cardiac troponin T in our cohort, thus, troponin was not included in our variable selection process.

The OURMAPCN score was specifically developed and optimized with the goal of identifying COVID-19 patients at a high risk of mortality at hospital admission. The score was sufficiently validated in multiple independent cohorts from China and Italy. The score shows robust performance with regard to predicting mortality across different populations, ethnicities and clinical practices. However, further studies with larger cohorts in different populations will be needed to establish whether the performance of the score is influenced by sample sizes or population heterogeneity.

Our data show that this score significantly outperformed two existing pneumonia scores, the MuLBSTA³⁸ and CURB-65 scores³⁹, which were initially developed to estimate the mortality due to pneumonia. Indeed, neither of these scores showed adequate predictive accuracy when applied to the same COVID-19 validation cohorts in this study. In addition, certain parameters required for the MuLBSTA score and CURB-65 score, such as bacterial infection, were not commonly available at hospital admission in COVID-19 patients, which hinders their application as a stratification and triage tool at admission. When compared to the COVID-GRAM score by Liang et al.²⁷, the two models had similar predictive ability in the population in Wuhan. However, the OURMAPCN score has better performance in patient populations outside of Wuhan city. Since five out of the ten risk factors in the COVID-GRAM score were patient-reported variables,

inadequate collection of these variables may lead to the underestimation of disease severity. This probably explained the decrease in the AUROC in the population outside of Wuhan city. Recently, a 4C Mortality Score was established for estimating the in-hospital mortality in patients from European countries⁹. Due to the Glasgow coma scale score was not routinely carried out for patients who admitted to hospitals in China, 4C Mortality Score was not able to be estimated in the population involved in this study.

A critical aspect of the OURMAPCN score is that its capability for predicting the risk of mortality is not significantly impacted by age, sex, or several commonly observed comorbidities that have been associated with COVID-19 severity^{1,40–44}. We recommended an OURMAPCN score of 11 as a cut-off threshold to maximize sensitivity while allowing some false positives.

The overall validity of the OURMAPCN score for the prediction of mortality was strengthened by the large cohort size, including in the training and testing cohorts; validation in multiple independent validation cohorts in China and Europe, and the well-balanced distribution in terms of age, sex and prevalence of comorbidities. Retrospective analyses indicated that only approximately 10–20% of patients with COVID-19 developed severe complications and need hospitalization or admission to the ICU, while the remaining patients are at lower risk of severe disease and mortality. If most patients at lower risk can recover from COVID-19 in community hospitals or isolation facilities, more medical resources can be allocated to patients at higher risk who need advanced therapy. Therefore, our scoring system provides a method of identifying patients with COVID-19 who are at high risk for severe complications and mortality. When medical resources are limited, prioritizing patients who are at high risk by allocating them more advanced medical resources is an effective means of reducing the mortality rate.

Limitation

Our study also has some limitations. *First*, our study was only based on baseline parameters at admission and focused on prediction. Hence no causal conclusions can be drawn from our algorithm^{45,46}. *Second*, the OURMAPCN score may simplify model utility and implementation with the resultant potential loss of information; however, the score performed well compared to the logistic model with eight-variables. *Third*, during the COVID-19 pandemic, several possible variables, including prehospital medication, personal habits, and socioeconomic factors, were not collected or were insufficiently documented. Even though rigorous validation of the data imputation process was conducted, the impacts of these unmeasured parameters and missing values on patient outcomes may reduce the power and accuracy of our model, especially when it is applied in more heterogeneous populations. *Fourth*, the variable selection process started from 64 candidate variables; however, the random forest model and LASSO model were set up to robustly limit the overfitting. Importantly, the developed score demonstrated good discrimination and calibration in all three validation cohorts.

Fifth, fewer patients in the study cohort were recruited in the very early period of the pandemic outbreak, which may create unknown bias in the temporal pattern of mortality following infection, such as delayed therapy, or the lack of knowledge and therapeutic experiences with COVID-19. *Sixth*, the patient cohort included COVID-19 in-hospital patients mainly from China, only one city in Italy. The generalizability of this model to patients of different genetic backgrounds and from different geographic environments was not examined. Extrapolation of the model to general or community patients with COVID-19 requires further examination. *Seventh*, the cutoff value used for dichotomization was derived according to the reference ranges of individual hospitals. Due to the stable performance of the score across the 22 hospitals involved in our study, we propose that other hospitals may use their own reference threshold for the cut-off value. Accumulating data from different institutions will provide us with increasing confidence in this conclusion in the future.

Conclusion

We have developed and validated a risk assessment tool, the OURMAPCN score, to evaluate the mortality risk in patients with COVID-19 using limited clinical parameters obtained at hospital admission. Previous data revealed that only approximately 10–20% of COVID-19 patients develop severe complications and mortality and therefore need hospitalization or admission to the ICU^{2,47}. Therefore, the OURMAPCN score may assist frontline clinicians in optimizing medical treatment in the context of limited resources by prioritizing patients who are at higher risk of mortality. While this risk score was derived from a patient population in China, it was validated in an independent European cohort. Satisfactory performance of the OURMAPCN score suggested that it could be a candidate tool to predict the mortality risk in COVID-19 patients elsewhere^{6,48–50}. However, before its clinical application, additional modification and validation of the OURMAPCN score to fit each country's specific situation are still needed.

Transparency

Declaration of funding

This work was supported by grants from the Special Foundation for Emergency Research on Prevention and Control of COVID-19 of Guangdong Province (2020B1111330003), National Key R&D Program of China (2016YFF0101504, 2019YFC2004700, 2020YFC0845500), the National Science Foundation of China (81630011, 81970364, 81970070, 81970011, 81870171, and 81570412), the Hubei Science and Technology Support Project (2019BFC582, 2018BEC473, and 2017BEC001), Medical flight plan of Wuhan University, and COVID-19 funds of Lombardia Region, CARIPO and Fondazione Umberto Veronesi to GC.

Declaration of financial/other relationships

The authors have no relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This

includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties. Peer reviewers on this manuscript have no relevant financial or other relationships to disclose.

Author contributions

ZC, JC, JZ, and FL designed study, collected and analyzed data, and wrote manuscript. FZ, PZ, JJQ, LHZ, YML, HTW, MMC, YCZ, LJS, XHS, XYZ, JZ, CZY, WFL, YQY, MYL, WMM, LML, PY, BX, PCL, ZXZ, ZGL, JHW, HFL, XGX, DHW, XFL, GP, LL, JY, GHC, BHZ, XDH and YFY collected and reviewed clinical, laboratory, and radiological data. JC, YML performed statistical analysis. JX, XW, JX, XZ, ZGS and DLG reviewed, interpreted, and checked clinical data. JC, ZGS, PZ and XJZ wrote manuscript and provided valuable suggestions for study design and data analysis. MC, EA, AA, GGS and GC collected, reviewed and analyzed clinical data in Italy and provided valuable suggestions for study design. JG, YBW, PZ, and H.L. contributed equally, designed the project, edited manuscript, and supervised the study. All authors have approved the final version of this paper.

Acknowledgements

None reported.

Data availability statement

The raw data supporting the findings of this study will be available from the corresponding authors after publication. The research team will provide an email address for communication once the data are approved to be shared with others. The proposal with detailed aims, statistical plan, and other information/materials might be required to guarantee the rationality of requirement and the security of data. The data without patient name and identifiers could be shared after reviewing and approving proposal and related materials.

Ethical approval

Studies in all cohorts were approved by the appropriate institutional review boards. Patient informed consent was waived by the ethics committees from each hospital.

Code availability

Codes will be available upon request to the corresponding author after publication.

Transparency statement

The corresponding author affirms that the manuscript is an honest, accurate, and transparent account of the study being reported. No important aspects of the study have been omitted, and all discrepancies from the study as planned have been explained.

References

- [1] Zhang P, Zhu L, Cai J, et al. Association of inpatient use of angiotensin converting enzyme inhibitors and angiotensin II receptor blockers with mortality among patients with hypertension hospitalized with COVID-19. *Circ Res.* 2020;126(12):1671–1681.
- [2] Emanuel EJ, Persad G, Upshur R, et al. Fair allocation of scarce medical resources in the time of COVID-19. *N Engl J Med.* 2020; 382(21):2049–2055.
- [3] World Health Organization. Coronavirus disease (COVID-19) pandemic. Geneva: WHO; 2020.
- [4] Liu H, Chen J, Yang Q, et al. Development and validation of a risk score using complete blood count to predict in-hospital mortality in COVID-19 patients. *Med.* 2021. DOI:10.1016/j.medj.2020.12.013
- [5] Tan L, Wang Q, Zhang D, et al. Lymphopenia predicts disease severity of COVID-19: a descriptive and predictive study. *Signal Transduct Target Ther.* 2020;5(1):33.
- [6] Bajwah S, Wilcock A, Towers R, et al. Managing the supportive care needs of those affected by COVID-19. *Eur Respir J.* 2020; 55(4):2000815.
- [7] Cai J, Li H, Zhang C, et al. The neutrophil-to-lymphocyte ratio determines clinical efficacy of corticosteroid therapy in patients with COVID-19. *Cell Metab.* 2021;33(2):258–269.e3.
- [8] Wynants L, Van Calster B, Bonten MMJ, et al. Prediction models for diagnosis and prognosis of COVID-19 infection: systematic review and critical appraisal. *BMJ.* 2020;369:m1328.
- [9] 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.
- [10] National Health Commission of China. New Coronavirus Pneumonia Prevention and Control Program. National Health Commission of China. 2020. <http://www.nhc.gov.cn>.
- [11] World Health Organization. Laboratory testing for 2019 novel coronavirus (2019-nCoV) in suspected human cases Interim guidance. Geneva: World Health Organization; 2020.
- [12] Stekhoven DJ, Buhlmann P. MissForest-non-parametric missing value imputation for mixed-type data. *Bioinformatics.* 2012;28(1): 112–118.
- [13] Bagherzadeh-Khiabani F, Ramezankhani A, Azizi F, et al. A tutorial on variable selection for clinical prediction models: feature selection methods in data mining could improve the results. *J Clin Epidemiol.* 2016;71:76–85.
- [14] Hastie T, Tibshirani R, Friedman J. Data mining, inference, and prediction. 2nd ed. Springer Series in Statistics. New York: Springer-Verlag; 2009.
- [15] Liu X, Song M, Tao D, et al. Random forest construction with robust semisupervised node splitting. *IEEE Trans Image Process.* 2015;24(1):471–483.
- [16] Fox EW, Hill RA, Leibowitz SG, et al. Assessing the accuracy and stability of variable selection methods for random forest modeling in ecology. *Environ Monit Assess.* 2017;189(7):316.
- [17] Lin H, Long E, Ding X, et al. Prediction of myopia development among Chinese school-aged children using refraction data from electronic medical records: a retrospective, multicentre machine learning study. *PLoS Med.* 2018;15(11):e1002674.
- [18] Bramer M. Pre-pruning classification trees to reduce overfitting in noisy domains. Berlin, Heidelberg: Springer; 2002.
- [19] Fürnkranz J. Pruning algorithms for rule learning. *Machine Learning.* 1997;27(2):139–172.
- [20] Yatsunenkov T, Rey FE, Manary MJ, et al. Human gut microbiome viewed across age and geography. *Nature.* 2012;486(7402): 222–227.
- [21] Mogensen UB, Ishwaran H, Gerds TA. Evaluating random forests for survival analysis using prediction error curves. *J Stat Softw.* 2012;50(11):1–23.
- [22] Epifanio I. Intervention in prediction measure: a new approach to assessing variable importance for random forests. *BMC Bioinformatics.* 2017;18(1):230.
- [23] Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. *J Stat Softw.* 2010; 33(1):1–22.
- [24] Stekhoven DJ. missForest: nonparametric missing value imputation using random forest [Internet]. 2013 [cited 2020 Nov 24]. Available from: <https://cran.r-project.org/web/packages/missForest/index.html>
- [25] Waljee AK, Mukherjee A, Singal AG, et al. Comparison of imputation methods for missing laboratory data in medicine. *BMJ Open.* 2013;3(8):e002847.

- [26] Cozzi L, Dinapoli N, Fogliata A, et al. Radiomics based analysis to predict local control and survival in hepatocellular carcinoma patients treated with volumetric modulated arc therapy. *BMC Cancer*. 2017;17(1):829.
- [27] Liang W, Liang H, Ou L, et al. Development and validation of a clinical risk score to predict the occurrence of critical illness in hospitalized patients with COVID-19. *JAMA Intern Med*. 2020;180(8):1081.
- [28] COVID-19 Treatment Guidelines Panel. Coronavirus disease 2019 (COVID-19) treatment guidelines. National Institutes of Health [Internet]. 2020 [cited 2020 Nov 24]. Available from: <https://www.covid19treatmentguidelines.nih.gov/>
- [29] Al-Gwaiz LA, Babay HH. The diagnostic value of absolute neutrophil count, band count and morphologic changes of neutrophils in predicting bacterial infections. *Med Princ Pract*. 2007;16(5):344–347.
- [30] Pepys MB, Hirschfield GM. C-reactive protein: a critical update. *J Clin Invest*. 2003;111(12):1805–1812.
- [31] Christ-Crain M, Jaccard-Stolz D, Bingisser R, et al. Effect of procalcitonin-guided treatment on antibiotic use and outcome in lower respiratory tract infections: cluster-randomised, single-blinded intervention trial. *Lancet*. 2004;363(9409):600–607.
- [32] Yang X, Jin Y, Li R, et al. Prevalence and impact of acute renal impairment on COVID-19: a systematic review and meta-analysis. *Crit Care*. 2020;24(1):356.
- [33] Liu YM, Xie J, Chen MM, et al. Kidney function indicators predict adverse outcomes of COVID-19. *Med*. 2021;2(1):38–48.e2.
- [34] 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(10229):1054–1062.
- [35] Qin JJ, Cheng X, Zhou F, et al. Redefining cardiac biomarkers in predicting mortality of inpatients with COVID-19. *Hypertension*. 2020;76(4):1104–1112.
- [36] Pranata R, Lim MA, Yonas E, et al. Body mass index and outcome in patients with COVID-19: a dose-response meta-analysis. *Diab Metab*. 2021;47(2):101178.
- [37] Gao YD, Ding M, Dong X, et al. Risk factors for severe and critically ill COVID-19 patients: a review. *Allergy*. 2021;76(2):428–455.
- [38] 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.
- [39] 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(5):377–382.
- [40] Zhu L, She ZG, Cheng X, et al. Association of blood glucose control and outcomes in patients with COVID-19 and pre-existing type 2 diabetes. *Cell Metab*. 2020;31(6):1068–1077 e3.
- [41] Zhang XJ, Qin JJ, Cheng X, et al. In-hospital use of statins is associated with a reduced risk of mortality among individuals with COVID-19. *Cell Metab*. 2020;32(2):176–187 e4.
- [42] Cheng X, Liu YM, Li H, et al. Metformin is associated with higher incidence of acidosis, but not mortality, in individuals with COVID-19 and pre-existing type 2 diabetes. *Cell Metab*. 2020;32(4):537–547 e3.
- [43] Liu PP, Blet A, Smyth D, et al. The science underlying COVID-19: implications for the cardiovascular system. *Circulation*. 2020;142(1):68–78.
- [44] Zhou F, Liu YM, Xie J, et al. Comparative impacts of ACE (angiotensin-converting enzyme) inhibitors versus angiotensin II receptor blockers on the risk of COVID-19 mortality. *Hypertension*. 2020;76(2):e15–e7.
- [45] Guan WJ, Ni ZY, Hu Y, et al. Clinical characteristics of coronavirus disease 2019 in China. *N Engl J Med*. 2020;382(18):1708–1720.
- [46] Lauer SA, Grantz KH, Bi Q, et al. The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: estimation and application. *Ann Intern Med*. 2020;172(9):577–582.
- [47] Wu Z, McGoogan JM. Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China: summary of a report of 72314 cases from the Chinese Center for Disease Control and Prevention. *JAMA*. 2020;323(13):1239.
- [48] Grasselli G, Zangrillo A, Zanella A, et al. Baseline characteristics and outcomes of 1591 patients infected with SARS-CoV-2 admitted to ICUs of the Lombardy Region, Italy. *JAMA*. 2020;323(16):1574.
- [49] Simonnet A, Chetboun M, Poissy J, et al. High prevalence of obesity in severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) requiring invasive mechanical ventilation. *Obesity*. 2020;28(7):1195–1199.
- [50] Caly L, Druce J, Roberts J, et al. Isolation and rapid sharing of the 2019 novel coronavirus (SARS-CoV-2) from the first patient diagnosed with COVID-19 in Australia. *Med J Aust*. 2020;212(10):459–462.