medRxiv preprint doi: https://doi.org/10.1101/2020.10.21.20217380; this version posted October 25, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

Development of a predictive risk model for severe COVID-19 disease using population-based administrative data

Jiandong Zhou MSc [1], Sharen Lee [2], Xiansong Wang MSc [3], Yi Li MD [4], William KK Wu PhD FRCPath [3], Tong Liu MD PhD [5], Zhidong Cao PhD [6], Daniel Dajun Zeng PhD [6], Ian Chi Kei Wong PhD FRCPCH [7,8], Bernard Man Yung Cheung MB BChir PhD FRCP [9], Qingpeng Zhang PhD * [1], Gary Tse PhD FRCP FFPH * [4]

[1] School of Data Science, City University of Hong Kong, Hong Kong, Hong Kong SAR, China
[2] Cardiovascular Analytics Group, Laboratory of Cardiovascular Physiology, Li Ka Shing Institute of Health Sciences, Hong Kong, China
[3] Li Ka Shing Institute of Health Sciences, The Chinese University of Hong Kong, Hong Kong, China
[4] Department of Cardiothoracic Surgery, Wuhan Asia Heart Hospital Affiliated to Wuhan University of Science and Technology, Hubei, China.
[5] Tianjin Key Laboratory of Ionic-Molecular Function of Cardiovascular disease, Department of Cardiology, Tianjin Institute of Cardiology, Second Hospital of Tianjin Medical University, Tianjin 300211, China
[6] Institute of Automation, Chinese Academy of Sciences, Beijing, China.
[7] Department of Pharmacology and Pharmacy, University of Hong Kong, Pokfulam, Hong Kong, China
[8] Medicines Optimisation Research and Education (CMORE), UCL School of Pharmacy, London UK
[9] Department of Medicine, University of Hong Kong, Pokfulam, Hong Kong, China

Prof. Gary Tse, PhD, FRCP, FFPH
Tianjin Key Laboratory of Ionic-Molecular Function of Cardiovascular Disease, Department of Cardiology, Tianjin Institute of Cardiology, Second Hospital of Tianjin Medical University, Tianjin 300211, China
Email: gary.tse@doctors.org.uk

Prof. Qingpeng Zhang, PhD
School of Data Science, City University of Hong Kong,
Hong Kong, China
Email: qingpeng.zhang@cityu.edu.hk
Phone: +852 3442-4727

Word count: 1519 words

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2020.10.21.20217380; this version posted October 25, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

Abstract

Background: Recent studies have reported numerous significant predictors for adverse outcomes in COVID-19 disease. However, there have been few simple clinical risk score for prompt risk stratification. The objective is to develop a simple risk score for severe COVID-19 disease using territory-wide healthcare data based on simple clinical and laboratory variables.

Methods: Consecutive patients admitted to Hong Kong's public hospitals between 1st January and 22nd August 2020 diagnosed with COVID-19, as confirmed by RT-PCR, were included. The primary outcome was composite intensive care unit admission, need for intubation or death with follow-up until 8th September 2020.

Results: COVID-19 testing was performed in 237493 patients and 4445 patients (median age 44.8 years old, 95% CI: [28.9, 60.8]); 50% male) were tested positive. Of these, 212 patients (4.8%) met the primary outcome. A risk score including the following components was derived from Cox regression: gender, age, hypertension, stroke, diabetes mellitus, ischemic heart disease/heart failure, respiratory disease, renal disease, increases in neutrophil count, monocyte count, sodium, potassium, urea, alanine transaminase, alkaline phosphatase, high sensitive troponin-I, prothrombin time, activated partial thromboplastin time, D-dimer and C-reactive protein, as well as decreases in lymphocyte count, base excess and bicarbonate levels. The model based on test results taken on the day of admission demonstrated an excellent predictive value. Incorporation of test results on successive time points did not further improve risk prediction.

Conclusions: A simple clinical score accurately predicted severe COVID-19 disease, even without including symptoms, blood pressure or oxygen status on presentation, or chest radiograph results.

medRxiv preprint doi: https://doi.org/10.1101/2020.10.21.20217380; this version posted October 25, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

Introduction

The coronavirus disease 2019 has a wide clinical spectrum, with disease severities ranging from completely asymptomatic to the need for intubation and death. For example, those with existing cardiac problems are more likely to suffer from more severe disease life courses (Guo et al., 2020;Li et al., 2020;Shi et al., 2020;Wang et al., 2020). Aside from comorbidities, numerous risk factors such as high D-dimer (Yao et al., 2020), neutrophil (Huang et al., 2020), and liver damage (Cai et al., 2020) and deranged clotting (Tang et al., 2020) have been associated with disease severity. Such patients may benefit from early aggressive treatment. However, to date there are only a few easy-for-use risk models that can be used for early identification of such at-risk individuals in clinical practice (Knight et al., 2020;Liang et al., 2020). The aim of the study is to extend these previous findings and develop a predictive risk score based on demographic, comorbidity, medication record and laboratory data using territory-wide electronic health records, without clinical parameters or imaging results. We hypothesized that incorporation of test results on successive time points would improve risk prediction.

Methods

Study design and population

This study was approved by the Institutional Review Board of the University of Hong Kong/Hospital Authority Hong Kong West Cluster. This was a retrospective, territory-wide cohort study of patients undergoing COVID-19 RT-PCR testing between 1st January 2020 and

medRxiv preprint doi: https://doi.org/10.1101/2020.10.21.20217380; this version posted October 25, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

22nd August 2020 in Hong Kong. The patients were identified from the Clinical Data Analysis and Reporting System (CDARS), a territory-wide database that centralizes patient information from individual local hospitals to establish comprehensive medical data, including clinical characteristics, disease diagnosis, laboratory results, and drug treatment details. The system has been previously used by both our team and other teams in Hong Kong [7-9]. Patients demographics, prior comorbidities, hospitalization characteristics before admission due to COVID-19, medication prescriptions, laboratory examinations of complete blood counts, biochemical tests, diabetes mellitus tests, cardiac function tests, c-reactive protein, and blood gas tests were extracted. The list of ICD-9 codes for comorbidity identification, codes of intubation procedure, medication prescriptions are detailed in the **Supplementary Tables 1 to 3**.

Outcomes and statistical analysis

The primary outcome was a composite of need for intensive care admission, intubation or all-cause mortality. Mortality data were obtained from the Hong Kong Death Registry, a population-based official government registry with the registered death records of all Hong Kong citizens linked to CDARS. The need for ICU admission and intubation were extracted directly from CDARS. Descriptive statistics are used to summarize baseline clinical characteristics of all patients with COVID-19 and based on the occurrence of the primary outcome. Continuous variables were presented as median (95% confidence interval [CI] or interquartile range [IQR]) and categorical variables were presented as count (%). The Mann-

medRxiv preprint doi: https://doi.org/10.1101/2020.10.21.20217380; this version posted October 25, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

Whitney U test was used to compare continuous variables. The $\chi 2$ test with Yates' correction was used for 2×2 contingency data. Univariate Cox regression identifies significant mortality risk predictors, which were used as input of multivariate Cox regression. Hazard ratios (HRs) with corresponding 95% CIs and p values were reported. An easy-for-use predictive model was developed using the beta coefficients of the multivariate Cox regression. Successive laboratory tests at least 24 hours apart were used. All statistical tests were two-tailed and considered significant if p value<0.001. They were performed using RStudio software (Version: 1.1.456) and Python (Version: 3.6).

Results

Basic characteristics

A total of 4445 patients (median age 44.8 years old, 95% CI: [28.9, 60.8]); 50% male) were diagnosed with the COVID-19 infection between 1st January 2020 and 22nd August 2020 in Hong Kong public hospitals or their associated ambulatory/outpatient facilities (**Table 1**). On follow-up until 8th September 2020, a total of 212 patients (4.77%) met the primary outcome of need for intensive care admission or intubation, or death. The survival curve is presented in **Figure 1A.** Details on Cox regression analyses are shown in the **Supplementary Tables 4 and 5**.

Development of a clinical risk score and validation

medRxiv preprint doi: https://doi.org/10.1101/2020.10.21.20217380; this version posted October 25, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

For clinical practice, it is impractical to precisely input the values of all variables assessed from the different domains of the health records. We built a simple, easy-to-use model based on the number of abnormalities observed in each domain (**Table 2**). Using Harrell's C-index and the area under the receiver operator characteristic curve (AUC) as performance evaluation metrics, we compared the prediction strengths of different criteria for the clinical risk score using a five-fold cross validation approach. The first model required cut-off values of different variables at baseline, and laboratory examinations on for each successive 24 hours was compared to cut-off to determine whether the criterion was met at each time point (**Supplementary Table 6**). In the second model, the criterion was met if the value was abnormal by standard laboratory criteria, without using optimal cut-off values (**Supplementary Table 7**). In the third model, laboratory test results are compared to the criteria without cut-off values, to see if they were met on successive testing (**Supplementary Table 8**). For example, if a particular criterion is met on day 1, then they will automatically fulfill the criteria for subsequent days.

Finally, we calculated the clinical risk score of each COVID-19 patient based on the third model and plotted the distribution of the whole cohort (**Figure 1B**). Patients meeting the primary outcome (n=212) have significantly higher risk score (median: 19, 95% CI: 14-24, max: 37) than those who did not (median: 8, 95% CI: 5-11, max: 34) **(Supplementary Table 9)**, indicating the performance of the clinical risk score (**Supplementary Table 10**). Receiver operating characteristic curve (ROC) of predicting adverse composite outcome of COVID-19 patients with the dichotomized risk score cut-off is shown in **Figure 1C**, demonstrating an AUC of 0.91. Further we generated the survival curves stratified by the dichotomized risk score

medRxiv preprint doi: https://doi.org/10.1101/2020.10.21.20217380; this version posted October 25, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

in **Figure 1D,** where yellow and blue curves represent the survival analysis for patients with a clinical risk score is larger and smaller than the cut-off, respectively.

Discussion

In this study, we developed a simple clinical score to predict severe COVID-19 disease based on age, gender, medical comorbidities, medication records, and laboratory examination results. This achieved good predictions with a c-statistic of 0.91, even without the consideration of clinical parameters such as symptoms, blood pressure, oxygen status on presentation or chest radiograph results.

COVID-19 disease has placed significant pressures on healthcare systems worldwide. Early risk stratification may better direct the use of limited resources and allow clinicians to triage patients and make clinical decisions based on limited evidence objectively. For example, low-risk patients may require simple monitoring only, whilst patients that are likely to deteriorate may benefit from intensive drug treatment or intensive care. Currently, the availability of simple clinical risk scores for risk stratification is limited. The COVID-GRAM predicts development of critical illness, based on symptoms, radiograph results, clinical and laboratory details (Liang et al., 2020). Similarly, the 4C Mortality Score included eight variables readily available at initial hospital assessment: age, sex, number of comorbidities, respiratory rate, peripheral oxygen saturation, level of consciousness, urea level, and C-reactive protein (score range 0-21 points) (Knight et al., 2020). These scores produced moderately accurate predictions with C-index values of 0.86 and 0.61-0.76, respectively. Our simple and

medRxiv preprint doi: https://doi.org/10.1101/2020.10.21.20217380; this version posted October 25, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

easy-to-use model was based on comorbidity, drug, and laboratory data only, without needing clinical assessment details or the need of chest imaging. The model based on test results taken on the day of admission already demonstrated an excellent predictive value with a C-index of 0.87. Incorporation of test results on successive time points did not further improve risk prediction, indicating that initial data are sufficient to produce accurate predictions of severe disease.

Limitations

The major limitation of this study is that it is based on a single territory-wide cohort. The model should be externally validated using patient data from other different regions.

Conclusion

A simple clinical score based on only demographics, comorbidities, medication records and laboratory tests accurately predicted severe COVID-19 disease, even without including symptoms on presentation, blood pressure, oxygen status or chest radiograph results. The model based on test results taken on the day of admission showed an excellent predictive value. Incorporation of test results on successive time points did not further improve risk prediction.

Conflicts of Interest

None.

medRxiv preprint doi: https://doi.org/10.1101/2020.10.21.20217380; this version posted October 25, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

References

Cai, Q., Huang, D., Yu, H., Zhu, Z., Xia, Z., Su, Y., Li, Z., Zhou, G., Gou, J., Qu, J., Sun, Y., Liu, Y., He, Q., Chen, J., Liu, L., and Xu, L. (2020). COVID-19: Abnormal liver function tests. *J Hepatol* 73**,** 566-574.

Guo, T., Fan, Y., Chen, M., Wu, X., Zhang, L., He, T., Wang, H., Wan, J., Wang, X., and Lu, Z. (2020). Cardiovascular Implications of Fatal Outcomes of Patients With Coronavirus Disease 2019 (COVID-19). *JAMA Cardiol*.

Huang, C., Wang, Y., Li, X., Ren, L., Zhao, J., Hu, Y., Zhang, L., Fan, G., Xu, J., Gu, X., Cheng, Z., Yu, T., Xia, J., Wei, Y., Wu, W., Xie, X., Yin, W., Li, H., Liu, M., Xiao, Y., Gao, H., Guo, L., Xie, J., Wang, G., Jiang, R., Gao, Z., Jin, Q., Wang, J., and Cao, B. (2020). Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* 395**,** 497-506.

Knight, S.R., Ho, A., Pius, R., Buchan, I., Carson, G., Drake, T.M., Dunning, J., Fairfield, C.J., Gamble, C., Green, C.A., Gupta, R., Halpin, S., Hardwick, H.E., Holden, K.A., Horby, P.W., Jackson, C., Mclean, K.A., Merson, L., Nguyen-Van-Tam, J.S., Norman, L., Noursadeghi, M., Olliaro, P.L., Pritchard, M.G., Russell, C.D., Shaw, C.A., Sheikh, A., Solomon, T., Sudlow, C., Swann, O.V., Turtle, L.C., Openshaw, P.J., Baillie, J.K., Semple, M.G., Docherty, A.B., and Harrison, E.M. (2020). 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* 370**,** m3339.

Li, X., Guan, B., Su, T., Liu, W., Chen, M., Bin Waleed, K., Guan, X., Gary, T., and Zhu, Z. (2020). Impact of cardiovascular disease and cardiac injury on in-hospital mortality in patients with COVID-19: a systematic review and meta-analysis. *Heart*.

Liang, W., Liang, H., Ou, L., Chen, B., Chen, A., Li, C., Li, Y., Guan, W., Sang, L., Lu, J., Xu, Y., Chen, G., Guo, H., Guo, J., Chen, Z., Zhao, Y., Li, S., Zhang, N., Zhong, N., He, J., and Covid-19, F.T.C.M.T.E.G.F. (2020). Development and Validation of a Clinical Risk Score to Predict the Occurrence of Critical Illness in Hospitalized Patients With COVID-19. *JAMA Internal Medicine* 180**,** 1081-1089.

Shi, S., Qin, M., Shen, B., Cai, Y., Liu, T., Yang, F., Gong, W., Liu, X., Liang, J., Zhao, Q., Huang, H., Yang, B., and Huang, C. (2020). Association of Cardiac Injury With Mortality in Hospitalized Patients With COVID-19 in Wuhan, China. *JAMA Cardiol*.

Tang, N., Li, D., Wang, X., and Sun, Z. (2020). Abnormal coagulation parameters are associated with poor prognosis in patients with novel coronavirus pneumonia. *J Thromb Haemost* 18**,** 844-847.

Wang, Y., Roever, L., Tse, G., and Liu, T. (2020). 2019-Novel Coronavirus-Related Acute Cardiac Injury Cannot Be Ignored. *Curr Atheroscler Rep* 22**,** 14.

Yao, Y., Cao, J., Wang, Q., Shi, Q., Liu, K., Luo, Z., Chen, X., Chen, S., Yu, K., Huang, Z., and Hu, B. (2020). D-dimer as a biomarker for disease severity and mortality in COVID-19 patients: a case control study. *Journal of Intensive Care* 8**,** 49.

Table 1. Baseline clinical characteristics of patients with COVID-19.

Comparisons were made between patients meeting the primary outcome vs. those that did not. * for p≤ 0.05, ** for p ≤ 0.01, *** for p ≤ 0.001

Variable	All (n=4445) Median (IQR) Max/n (%)	Primary outcome (n=212)	P value
Demographics			
Male gender	2229(50.14%);	140(66.03%)	0.0023**
Age, year	44.8(28.9-60.8);100.6;n=4445	71.0(61.9-82.6);98.7	p<0.0001***
Comorbidities			
Cardiovascular	323(7.26%)	63(29.71%)	p<0.0001***
Respiratory	4319(97.16%)	207(97.64%)	0.4612
Kidney	239(5.37%)	70(33.01%)	p<0.0001***
Endocrine	39(0.87%)	6(2.83%)	0.0025**
Diabetes mellitus	507(11.40%)	77(36.32%)	p<0.0001***
Hypertension	704(15.83%)	117(55.18%)	p<0.0001***
Gastrointestinal	4291(96.53%)	202(95.28%)	0.7681
Stroke	191(4.29%)	46(21.69%)	p<0.0001***
Hospitalization			
No. of episodes	1.0(1.0-2.0);98.0;n=4284	1.0(1.0-3.0);13.0;n=193	0.6634
Length of Stay, days	13.0(8.0-19.0);351.0;n=4284	23.0(14.0-36.0);155.0;n=193	p<0.0001***
No. of Emergency Readmissions	1.0(1.0-2.0);32.0;n=189	2.0(1.0-3.0);12.0;n=35	0.003**
Medication prescriptions			
ACEI	251(5.64%)	56(26.41%)	p<0.0001***
ARB	242(5.44%)	30(14.15%)	p<0.0001***
Steroids	632(14.21%)	38(17.92%)	0.0512.
Kaletra	102(2.29%)	70(33.01%)	p<0.0001***
Ribavirin	1127(25.35%)	75(35.37%)	0.0021**
Interferon beta	1865(41.95%)	165(77.83%)	p<0.0001***

Hydroxychloroquine	92(2.06%)	12(5.66%)	0.0023**
Calcium channel blockers	755(16.98%)	117(55.18%)	p<0.0001***
Beta blockers	328(7.37%)	65(30.66%)	p<0.0001***
Diuretics for heart failure	193(4.34%)	102(48.11%)	p<0.0001***
Diuretics for hypertension	15(0.33%)	7(3.30%)	p<0.0001***
Nitrates	37(0.83%)	3(1.41%)	0.0125*
Antihypertensive drugs	150(3.37%)	35(16.50%)	0.0823.
Complete blood count tests			
MCV, fL	86.8(83.0-90.1);110.6;n=2939	88.7(84.8-92.35);105.4;n=170	0.0061.
Basophil, x10^9/L	0.01(0.01-0.02);0.2;n=3684	0.003(0.0-0.02);0.13;n=186	0.0123**
Eosinophil, x10^9/L	0.03(0.0-0.1);2.25;n=3839	0.0(0.0-0.01);1.25;n=201	0.1236
Lymphocyte, x10^9/L	1.4(1.0-1.86);16.99;n=3848	0.8(0.6-1.27);3.09;n=201	p<0.0001***
Metamyelocyte, x10^9/L	0.105(0.065-0.24);0.81;n=16	0.22(0.08-0.505);0.81;n=8	0.0012**
Monocyte, x10^9/L	0.5(0.37-0.63);3.26;n=3848	0.5(0.35-0.7);2.1;n=201	0.871
Neutrophil, x10^9/L	3.30(2.4-4.38);27.17;n=3848	4.68(3.43-7.1);27.17;n=201	p<0.0001***
WBC, x10^9/L	5.404(4.31-6.76);29.64;n=3918	6.18(4.9-8.57);29.64;n=205	0.0234*
MCH, g/dL	29.7(28.29-30.8);36.8;n=3918	30.0(28.5-31.3);36.8;n=205	0.2745
Myelocyte, x10^9/L	0.195(0.08-0.36);1.29;n=38	0.33(0.2-0.4);1.29;n=18	p<0.0001***
Platelet, x10^9/L	217.0(177.0-272.0);778.0;n=3917	183.0(139.0-230.0);778.0;n=204	p<0.0001***
Reticulocyte, x10^9/L	45.55(30.45-66.95);318.0;n=36	42.8(33.15-51.35);57.0;n=4	0.0076**
HCT, L/L	0.40(0.37-0.431);0.53;n=3385	0.38(0.35-0.42);0.51;n=165	0.1835
RBC, x10^12/L	4.68(4.33-5.08);7.45;n=3919	4.38(4.02-4.77);6.79;n=205	0.7623
Biochemical tests			
Potassium, mmol/L	3.8(3.57-4.1);5.9;n=2995	3.9(3.57-4.24);5.9;n=190	0.0124*
Urate, mmol/L	0.30(0.243-0.39);0.64;n=127	0.34(0.30-0.47);0.62;n=18	0.2341
Albumin, g/L	40.6(37.0-44.0);146.0;n=3001	35.0(29.25-38.95);146.0;n=190	p<0.0001***
Sodium, mmol/L	139.0(137.0-140.7);149.4;n=3002	137.0(134.2-141.0);149.42;n=190	0.1246

Urea, mmol/L	3.9(3.1-4.94);41.95;n=3000	6.4(4.52-8.85);41.95;n=190	p<0.0001***
Protein, g/L	74.0(70.0-77.5);93.2;n=2683	70.0(65.1-75.6);93.2;n=164	0.0523*
Creatinine, umol/L	70.0(59.0-84.0);1280.0;n=3002	84.0(69.5-110.9);1280.0;n=190	p<0.0001***
Alkaline Phosphatase, U/L	65.0(54.0-80.0);932.0;n=2988	69.5(57.0-93.0);932.0;n=189	0.0018**
Aspartate Transaminase, U/L	27.0(21.0-40.0);1807.4;n=992	41.0(27.0-66.5);1807.4;n=86	p<0.0001***
Alanine Transaminase, U/L	23.0(16.0-35.0);583.0;n=2426	30.7(20.0-50.0);380.0;n=148	p<0.0001***
Bilirubin, µmol/L	7.8(5.6-11.0);235.0;n=2987	9.2(6.2-14.0);235.0;n=189	0.0035**
Glucose, mmol/L	5.6(5.03-6.8);29.29;n=2945	7.2(5.8-9.4);29.29;n=187	0.1835
Diabetes mellitus tests			
HbA1C, mmol/mol	13.4(12.1-14.6);18.07;n=3919	12.6(9.4-14.2);17.2;n=205	0.5334
Cholesterol	3.87(2.72-4.8);9.72;n=663	3.69(2.60-4.53);7.2;n=107	0.3451
D-dimer, ng/mL	293.3(165.0-565.9);11200.0;n=805	834.0(364.0-1293.0);11200.0;n=73	p<0.0001***
High sensitive troponin-I, ng/L	3.0(1.1-7.1);3876.28;n=1231	10.0(4.6-32.9);652.28;n=108	p<0.0001***
Lactate dehydrogenase, U/L	196.0(165.9-242.0);1374.0;n=3156	321.0(230.6-437.5);1116.0;n=179	p<0.0001***
Prothrombin Time/INR, second	12.0(11.5-12.6);43.4;n=1455	12.4(11.7-13.15);26.8;n=132	0.5313
APTT, second	31.0(27.7-34.5);120.0;n=2053	33.0(28.4-36.0);120.0;n=187	0.3245
C-Reactive Protein, mg/dL	0.35(0.12-1.31);34.0;n=3854	4.57(1.4-10.31);34.0;n=207	p<0.0001***
Blood Gas Tests			
HCO3/Bicarbonate, mmol/L	23.3(20.2-26.2);33.2;n=423	22.9(20.5-25.3);32.8;n=68	0.0325*
Base Excess, mmol/L	-0.5(-2.55-1.55);9.5;n=512	-1.4(-3.75-0.65);6.4;n=180	p<0.0001***
Blood pCO2, kPa	4.88(4.15-5.7);10.94;n=513	4.6(3.9-5.5);10.91;n=180	0.6623
Blood PO2, kPa	11.68(9.5,14,64);79.4;n=321	11.23(9.2,14.23);79.4;n=180	0.4512
Blood pH	7.4(7.4-7.5);7.61;n=513	7.43(7.37-7.48);7.6;n=180	0.0921.
Calcium, mmol/L	1.13(1.09-1.175);1.33;n=36	1.13(1.07-1.19);1.33;n=25	0.2341

medRxiv preprint doi: https://doi.org/10.1101/2020.10.21.20217380; this version posted October 25, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

Table 2. Easy-to-use score system for early prediction of severe COVID-19 disease.

Risk factor	**Cut-off**	**Score**
Male gender	Male	4
Age	60-64 years old	2
	65-69 years old	3
	70-74 years old	3
	≥75 years old	10
Hypertension or on any anti-hypertensive drug	Present	1
Stroke	Present	2
Diabetes mellitus	Present	1
Myocardial infarction, heart failure or on any medication for these conditions	Present	4
Respiratory disease	Present	3
Kidney disease	Present	2
High neutrophil	5.44 x10^9/L	1
High monocyte	0.80 x10^9/L	1
Low lymphocyte	0.77 x10^9/L	2
High sodium	136.4 mmol/L	2
High potassium	4.32 mmol/L	1
High urea	20.35 mmol/L	5
High ALT	56 U/L	1
High ALP	143 IU/L	1
High sensitive troponin-I, ng/L	7.6 ng/L	1
High prothrombin time/INR	12.6 seconds	1
High APTT	38.8 seconds	2
High D-dimer	831.46 ng/mL	1
High C-reactive protein	0.88 mg/dL	2
Low base excess	-3.5	3
Low bicarbonate	28.3	2

medRxiv preprint doi: https://doi.org/10.1101/2020.10.21.20217380; this version posted October 25, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

Figures

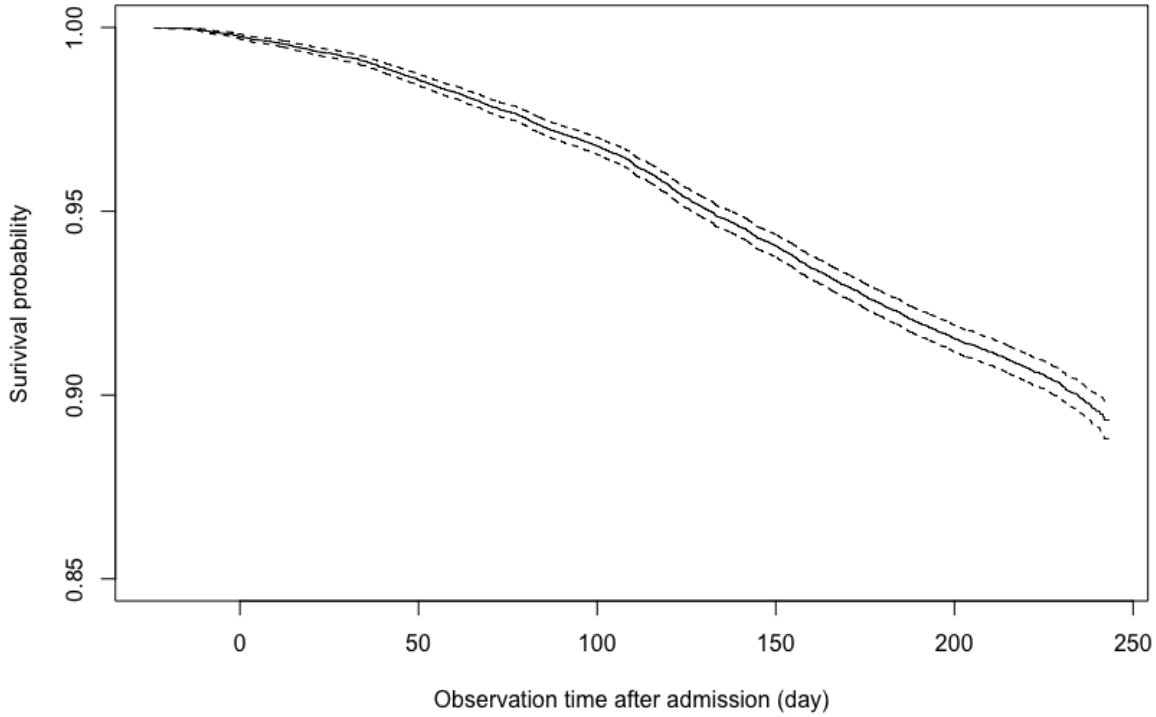

Figure 1A. Survival curve of COVID-19 patient for the primary outcome, a composite of intensive care admission, need for intubation or death.

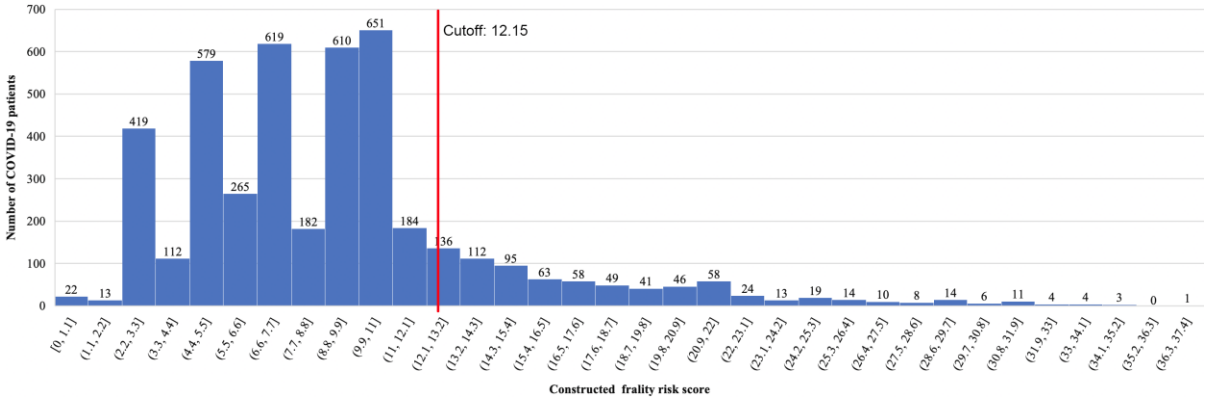

Figure 1B. Distribution of derived risk score for composite outcome identification.

medRxiv preprint doi: https://doi.org/10.1101/2020.10.21.20217380; this version posted October 25, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

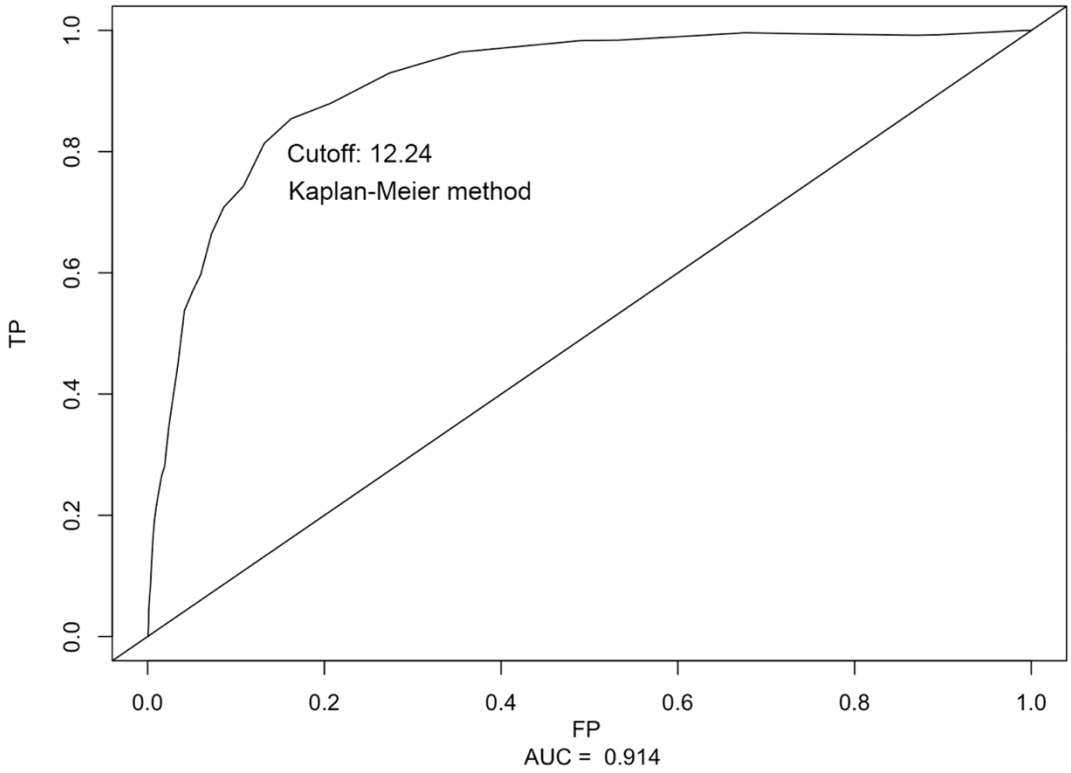

Figure 1C. Receiver operating characteristic curve (ROC) of predicting adverse composite outcome of COVID-19 patients with dichotomized risk score.

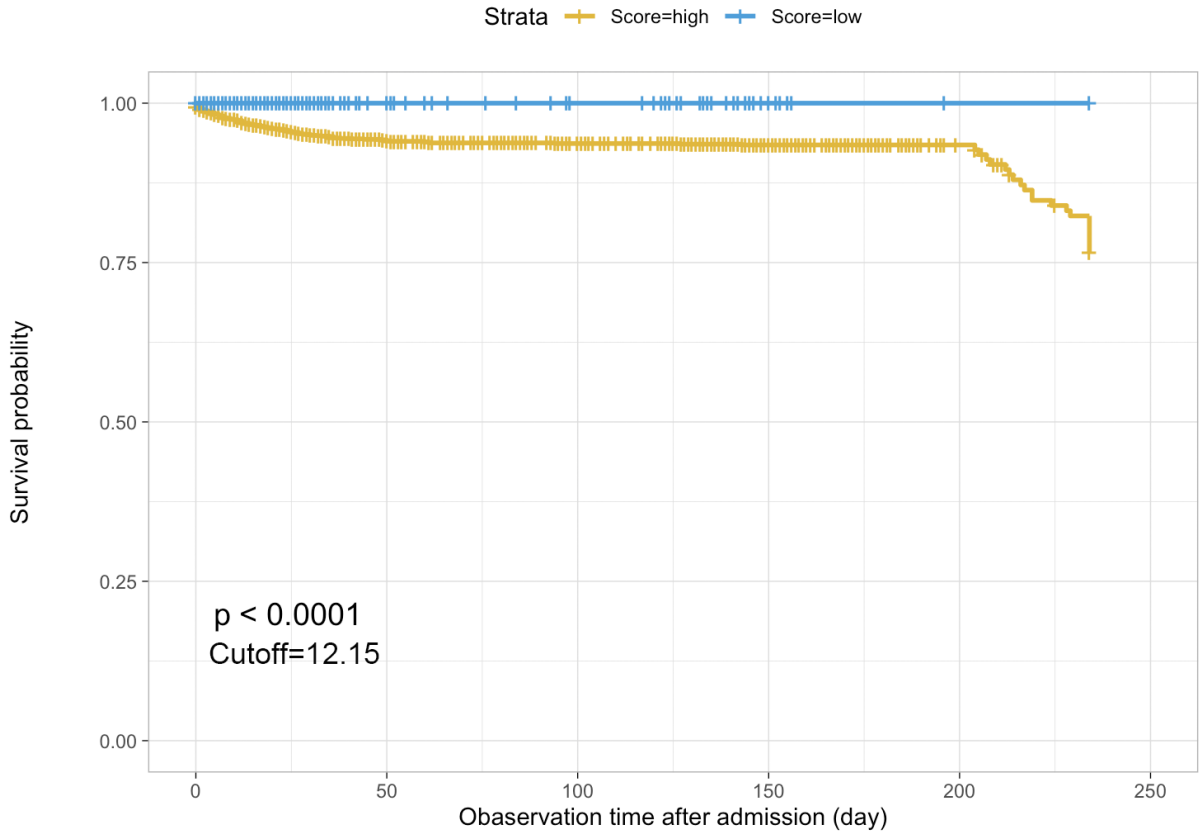

Figure 1D. Survival curve of COVID-19 patients stratified by dichotomized risk score.