EUROPEAN RESPIRATORY *journal*

FLAGSHIP SCIENTIFIC JOURNAL OF ERS

Early View

Research letter

A predictive model for disease progression in non-severe illness patients with Corona Virus Disease 2019

Mengyao Ji, Lei Yuan, Wei Shen, Junwei Lv, Yong Li, Jia Chen, Chaonan Zhu, Bo Liu, Zhenzhen Liang, Qiang Lin, Wenjie Xie, Ming Li, Zhifan Chen, Xuefang Lu, YiJuan Ding, Ping An, Sheng Zhu, Mengting Gao, Hao Ni, Lanhua Hu, Guanglei Shi, Lei Shi, Weiguo Dong

Please cite this article as: Ji M, Yuan L, Shen W, *et al*. A predictive model for disease progression in non-severe illness patients with Corona Virus Disease 2019. *Eur Respir J* 2020; in press (https://doi.org/10.1183/13993003.01234-2020).

This manuscript has recently been accepted for publication in the *European Respiratory Journal*. It is published here in its accepted form prior to copyediting and typesetting by our production team. After these production processes are complete and the authors have approved the resulting proofs, the article will move to the latest issue of the ERJ online.

A predictive model for disease progression in non-severe illness patients with Corona Virus Disease 2019

Mengyao Ji[*,1], Lei Yuan[*,2], Wei Shen[*,3,4], Junwei Lv[3,4], Yong Li[*,5], Jia Chen[3,4], Chaonan Zhu[3,4], Bo Liu[3,4], Zhenzhen Liang[6], Qiang Lin[3,4], Wenjie Xie[7], Ming Li[2], Zhifan Chen[8], Xuefang Lu[9], YiJuan Ding[1], Ping An[1], Sheng Zhu[3,4], Mengting Gao[2], Hao Ni[3,4], Lanhua Hu[9], Guanglei Shi[3,4], Lei Shi[Δ,3,4], Weiguo Dong[Δ,1]

[1] Department of Gastroenterology, Renmin Hospital of Wuhan University, Wuhan, Hubei, China;
[2] Department of Information Center, Renmin Hospital of Wuhan University, Wuhan Hubei, China;
[3] Shanghai Key Laboratory of Artificial Intelligence for Medical Image and Knowledge Graph;
[4] Hangzhou YITU Healthcare Technology Co., Ltd.
[5] Department of Radiology, Suining Central Hospital, Suining, Sichuan, China;
[6] Zhejiang Provincial Center for Disease Control and Prevention
[7] Department of Critic Care Medicine, Renmin Hospital of Wuhan University, Wuhan, P.R. China;
[8] Department of Radiology, Affiliated Hospital of Chengdu University, Chengdu, Sichuan, China;
[9] Department of Radiology, Renmin Hospital of Wuhan University, Wuhan, Hubei, China;
[*] Mengyao Ji, Lei Yuan, Wei Shen and Yong Li contributed equally to this paper.
[Δ] Correspondence: Weiguo Dong; weiguodong_66@163.com; Lei Shi; lei.shi@yitu-inc.com

Backgrounds

For nearly three months, COVID-19 broke out across China and spread around the world [1]. This disease caused varying degrees of illness. The proportion of patients with COVID-19 in the non-severe illness was 84.3% on admission, and severe cases were account for 15.7%[2]. Most of the non-severe pneumonia patients would gradually alleviate and cure during treatment, while others would rapidly progress to severe illness, which had a poor prognosis [3-4]. As recent reports, the cumulative risk of the composite end point was 3.6% in all the COVID-19 patients, and the cumulative risk was 20.6%[2] for the severe illness.

Although it was still unknown whether early identification and intervention for non-severe patients with COVID-19 could prevent progression into severe degree. According to the experience of treating other diseases, there might be a great promoting effect for the treatment. In this paper, we aim to build a predictive model for identifying high-risk non-severe pneumonia patients at early stage.

Methods

Eighty-six patients with COVID-19 in non-severe pneumonia on admission were recruited as training cohort at Renmin Hospital of Wuhan University from January 2nd to 20th, 2020, and another 62 patients were prospectively enrolled as validation cohort from January 28th to February 9th, 2020. COVID-19 was confirmed by real-time PCR. Disease severities of COVID-19 were defined as severe and non-severe pneumonia based on the criteria of American Thoracic Society guidelines for community-acquired pneumonia [2, 5]. The exclusion criteria included: (1) degrees of severity were not available on admission or during follow-up; (2) diagnosed with severe illness at the time of admission; (3) confirmed with COVID-19 and treated at other hospitals; (4) medication was administered within 15 days before admission; (5) received oxygen support during follow-up. Patients were divided into progressed or non-progressed group based on whether they progressed to severe illness or not during the 14-day follow-up period. Comorbidity included diabetes, hypertension, cardiovascular and cerebrovascular diseases, chronic obstructive pulmonary disease, malignant tumor, chronic liver disease, chronic kidney disease, tuberculosis and immunodeficiency diseases, etc.

Clinical characteristics and laboratory findings were extracted from electronic medical records. The radiological features were extracted from the chest CT imaging using a double-blind method[6]. To evaluate the lesion size accurately, a diagnosis system for COVID-19 based on artificial intelligence (AI) was employed to measure volume ratio of pneumonia automatically by analyzing CT value [7-8].

Logistic regression was used as the classifier to build the predictive model. The discrimination performance of the predictive model was quantified by the value of area under the ROC curve (AUC) in the cross-validation of training and validation dataset. Risk index calculated with the weight of each variable in the model was used to identify high-risk groups. All analyses were performed using R-3.6.0.

Results

The median age of the 148 patients was 46.5 years (IQR 35.8–58.0), and 81 (54.7%) were female. A total of 60 (40.5%) non-severe patients progressed to severe illness, and the median time of progression was 5.0 days (IQR 2.8-9.0). For training cohort, 60 (40.5%) non-severe patients progressed to severe illness, and 26 (41.9%) cases were in validation cohort. The median days of progression in these two cohorts were 5.5 days (IQR 1.0-9.0) and 5.0 days (IQR 3.0-9.8). Description of variables was provided in the Table 1.

To build the predictive model, we tested all the clinical, laboratory and radiological variables, except for the characteristics about treatment. Four variables were finally included in the model, including comorbidity (β=1.234, P=0.036), dyspnea on admission (β=1.583, P=0.095), lactate dehydrogenase (β=0.007, P=0.027) and lymphocyte count (β=-2.012, P=0.002). The Hosmer Lemeshow test of the training dataset was done (χ^2=10.451, P=0.235). The AUC value in the cross-validation of training dataset was 0.819 (95% CI: 0.731-0.907). It was 0.759 (95% CI: 0.635-0.884) in the validation dataset. According to the regression coefficients, we gave the four variables different weights. The comorbidity was 12 points per unit, dyspnea was 16, lactate dehydrogenase was 0.07, and lymphocyte count was -20. Then, total scores for each person were calculated, and different scores showed different risks. AUC value based on the risk scores in training dataset was 0.856 (95% CI: 0.776-0.935). Patients were divided into high-risk group and low-risk group (total score >-6.0 and ≤-6.0)

based on the best cutoff value determined by the Youden index, and the sensitivity was 0.941, specificity was 0.635, More details could be found in Table 1.

Table 1 Description of clinical characteristics and multivariate analysis in training cohort

Variables	Clinical Characteristics		Multivariate analysis in training cohort	
	Training cohort (n=86)	Validation cohort (n=62)	OR (95% CI)	Score
Group				
Non-progressed group	52(60.5%)	36(58.1%)		
Progressed group	34(39.5%)	26(41.9%)		
Time of progression, days	5.5 (1.0-9.0)	5.0 (3.0-9.8)		
Age, years	50.5(37.0-60.5)	44.5(35.0-53.0)		
Age range, years				
<40	27(31.4%)	21(33.9%)		
40-49	15(17.4%)	14(22.6%)		
50-59	22(25.6%)	15(24.2%)		
60-69	13(15.1%)	9(14.5%)		
70-79	7(8.1%)	2(3.2%)		
⩾80	2(2.3%)	1(1.6%)		
Female	45(52.3%)	36(58.1%)		
Comorbidity	42(48.8%)	15(24.2%)	3.436(1.084-10.896)	12× (0/1; no = 0, yes = 1)
Dyspnea on admission	11(12.8%)	6(9.7%)	4.869(0.760-31.212)	16× (0/1; no = 0, yes = 1)
Temperature on admission, °C	36.8(36.5-37.2)	36.8(36.5-37.1)		
Respiratory rate on admission	19.0(18.0-20.0)	20.0(19.0-20.0)		
Lactate dehydrogenase, U/L	214.0(187.8-275.8)	201.5(160.3-247.0)	1.008(1.001-1.014)	0.07×per unit (U/L)
Procalcitonin, ng/mL	0.04(0.03-0.07)	0.03(0.02-0.05)		
Lymphocyte count, 10^9 /L	1.2(0.9-1.6)	1.3(1.0-1.7)	0.134(0.038-0.471)	-20×per unit (10^9/L)
White blood cells, 10^9 /L	4.8(3.7-6.1)	4.7(4.0-6.1)		
Neutrophil count, 10^9/L	3.1(2.2-4.1)	3.0(2.0-3.9)		
Platelet count, 10^9/L	159.3(132.5-204.0)	164.5(120.3-210.4)		
Hemoglobin concentration, g/L	138.5(127.0-156.6)	143.3(130.0-152.8)		
Arterial oxygen saturation (%)	97.0(95.3-98.8)	96.0(95.0-98.0)		
Radiological abnormality				

GGOSS	36(41.9%)	-
Pure ground-glass opacity	32(37.2%)	-
Consolidation	12(14.0%)	-
Other	6(7.0%)	-
No. of affected segments	7.0(2.3-12.0)	-
Lesion size		
< 1cm	4(4.7%)	-
1-3cm	32(37.2%)	-
3cm-50% lobe	45(52.3%)	-
> 50% lobe	5(5.8%)	-
AI-based volume ratio of pneumonia		
-700~500	0.18(0.11-0.27)	-
-600~500	0.11(0.07-0.17)	-
Treatment		
Corticosteroid agents	55(64.0%)	19(30.6%)
Anti-infection agents	85(98.8%)	52(83.9%)
Interferon agents	34(39.5%)	7(11.3%)
Antiviral agents	74(86%)	61(98.4%)
Gamma globulin agents	54(62.8%)	21(33.9%)

Note：Data are presented as medians (interquartile ranges, IQR) and N (%). Percentages may not total 100 because of rounding. Variables in validation cohort were not completely collected, as some of them did not appear in the model of the training cohort. Abbreviations: GGOSS, ground-glass opacities overlapped with striped shadows

Discussions

In our prediction model, comorbidity was associated with disease progression, which meant that patients with comorbidities were more likely to progression than those without of them. Previous studies have shown that a higher proportion of patients with comorbidities in severe patients[9]. We further confirmed non-severe patients with comorbidities were easier to progress. It should be explained that the P value of dyspnea on admission was not less than 0.05 in the multivariate regression, which might be due to the relationship between dyspnea and the outcome in this study was not strictly linear after adjusting for other variables. Although we did try other models with better performance earlier, we finally chose logistic model because of its interpretability and simplicity of application. Patients progressed would be more likely to accompany with a decrease in lymphocyte count and an increase in lactate dehydrogenase [2, 10]. Our research further confirmed that these two indicators were also related to the progression. A decrease in lymphocyte count usually indicated the decline of immune function, and multiple organ dysfunction might lead to an increase in lactate dehydrogenase[11], which were consistent with the phenomenon we had observed clinically.

Previous reports have pointed out that advanced age was one of the risk factors for poor prognosis in patients with COVID-19[2, 3]. However, age was not included in the model. It suggests that treatment for young non-severe illness patients should not be neglected in early treatment. We speculated that age's contribution to disease progression was reflected in comorbidity and dyspnea. As well, some studies reported the correlations

between radiological indicators and COVID-19 disease [12]. Although the radiological features in CT images on admission were described in detail, they were not included into the model. We speculated that multiple images during the treatment instead of a single image could indicate the further progression of the disease. Although variables extracted with AI from CT imaging were not included in the model, this was a good attempt and would be the focus of our subsequent research.

There were some limitations in this study. First, patients with COVID-19 included in this study were from a single hospital, which would be a potential constraint for the generalization of our model. Second, critically ill patients were transferred to other designated hospitals according to the regulations of the local government. We were unable to track the patients' deaths in the short term, and the association between the model and overall survival could not be evaluated, which was a major regret in this study.

Conclusively, the progression of non-severe patients with COVID-19 could be predicted by our model based off clinical characteristics on admission. And the model was verified with a prospective validation cohort with well performance. With the help of our model, clinicians could easily identify high-risk non-severe patients on admission with few clinical indicators routinely, and thereby contributing to the treatment and prevention of COVID-19.

References

1. Chinese Center for Disease Control and Prevention. Distribution of Covid-19. Available from: http://2019ncov.chinacdc.cn/2019-nCoV/

2. WJ G, ZY N, Y H, WH L, CQ O, JX H, L L, H S, CL L, DSC H, B D, LJ L, G Z, KY Y, RC C, CL T, T W, PY C, J X, SY L, JL W, ZJ L, YX P, L W, Y L, YH H, P P, JM W, JY L, Z C, G L, ZJ Z, SQ Q, J L, CJ Y, SY Z, NS Z, medicine JTNEjo. Clinical Characteristics of Coronavirus Disease 2019 in China. 2020.

3. D W, B H, C H, F Z, X L, J Z, B W, H X, Z C, Y X, Y Z, Y L, X W, JAMA PZJ. Clinical Characteristics of 138 Hospitalized Patients With 2019 Novel Coronavirus-Infected Pneumonia in Wuhan, China. 2020.

4. X Y, Y Y, J X, H S, J X, H L, Y W, L Z, Z Y, M F, T Y, Y W, S P, X Z, S Y, medicine SYJTLR. Clinical course and outcomes of critically ill patients with SARS-CoV-2 pneumonia in Wuhan, China: a single-centered, retrospective, observational study. 2020.

5. JP M, GW W, AC L, A A, J B, K C, LA C, NC D, MJ F, SA F, MR G, ML M, DM M, MI R, respiratory WCJAjo, medicine cc. Diagnosis and Treatment of Adults with Community-acquired Pneumonia. An Official Clinical Practice Guideline of the American Thoracic Society and Infectious Diseases Society of America. 2019: 200(7): e45-e67.

6. F S, N S, F S, Z Z, J S, H L, Y L, Y J, Radiology SYJ. Emerging Coronavirus 2019-nCoV Pneumonia. 2020: 200274.

7. JG C, JM G, RE Y, CM P, CH L, B vG, DH C, radiology KYJE. Software performance in segmenting ground-glass and solid components of subsolid nodules in pulmonary adenocarcinomas. 2016: 26(12): 4465-4474.

8. A K, Y K, S H, K S, S U, R G, T S, thoracic KMJG, surgery c. One-dimensional mean computed tomography value evaluation of ground-glass opacity on high-resolution images. 2012: 60(7): 425-430.

9. W L, W G, R C, W W, J L, K X, C L, Q A, W L, H L, S L, Oncology HJJTL. Cancer patients in SARS-CoV-2 infection: a nationwide analysis in China. 2020: 21(3): 335-337.

10. Y L, Y Y, C Z, F H, F W, J Y, Z W, J L, J L, C F, Z Z, L W, L P, L C, Y Q, D Z, S T, L Y, J X, C Z, C J, sciences LLJSCL. Clinical and biochemical indexes from 2019-nCoV infected patients linked to viral loads and lung injury. 2020: 63(3): 364-374.

11. AA K, KS A, FA A, SJT G, Rahmani AH %J Endocrine m, targets idd. The Biochemical and Clinical Perspectives of Lactate Dehydrogenase: An Enzyme of Active Metabolism. 2019.

12. YM A, AA A, HH B, H N, AS A, A G, H H, A A, A K, medicine ARBJAoi. Clinical course and outcomes of

critically ill patients with Middle East respiratory syndrome coronavirus infection. 2014: 160(6): 389-397.