Development and Validation of a Nomogram to Predict Deteriorating Trajectory in Patients with COVID-19 Infection: A Population-Based Prospective Study

Wei Chen
Hubei cancer hospital
Menglin Zhu
Enshi hospital
Jian Li
Enshi hospital
Cuiping Pan
Hubeicancer hospital
Demian Zhao
Hubei cancer hospital
Yuting Jin
Hubei cancer hospital
Manxiu Li
Hubei cancer hospital
Shun Wu
Hubei cancer hospital
Yaojun Feng
Hubei cancer hospital
Tiejun Wang (tiejunwanghp@163.com)
Hubei Cancer Hospital

Research

Keywords: COVID-19, mild-moderate patients, severe probability, nomogram score

Posted Date: June 2nd, 2020

DOI: https://doi.org/10.21203/rs.3.rs-30640/v1

Abstract

Background Most of the patients with COVID-19 infection are mild to moderate initially. However, there is no effective prediction for the patients to develop into severe or extremely severe. This study aims to develop an effective clinical prediction model.

Methods A single-center, retrospective, observational study conducted. A nomogram was conducted based on the results of multivariate logistic regression analysis.

Results A total of 483 patients diagnosed mild to moderate were included, among these patients 62 developed severe or extremely critical illness. Seven variables including hyperlipidemia, vomiting, diarrhea, lymphocyte, imaging and mentality were associated with deteriorating trajectory. The ROC curve showed that model was robust, for which the area under the curve of the training set and the validation set are 0.873 and 0.813.

Conclusions For patients with mild to moderate COVID-19 infection, nomogram score can effectively predict the possibility of patients developing into severe or extremely critical.

Background

Since December 2019, COVID-19 has shown an outbreak trend. This new coronavirus has swept the world since Wuhan, China. As of June 2020, the number of people worldwide who had confirmed the diagnosis of COVID-19 had exceeded one million[1, 2].Starting from the evolution of the patient's condition, the majority of patients with COVID-19 infection are mild and moderate, and only a small number of patients are severe and extremely critical. However, even if the patient is diagnosed as mild or moderate, which still needs to be effectively isolated, because COVID-19 has significant human-to-human characteristics[3].

The Chinese government quickly deployed and established a sheltered hospital for accommodating patients with mild to moderate COVID-19, and achieved remarkable results[4]. Therefore, it is very important to carry out early isolation of mild to moderate patients, and to predict and intervene in patients who may develop into severe and extremely severe diseases.

To the best of our knowledge, existing studies have reported risk factors for COVID-19 patients to develop severe or very critical illnesses, but few reports have analyzed and studied the development of mild to moderate patients to severe illnesses[3, 5, 6]. In particular, there is a lack of predictive analysis of mild to moderate patients who may develop into severe illness.

Collectively, prediction of risk factors seems to be a promising way for predicting severe probability after COVID-19 infection. Here, we conducted a prospective observational study based on the nomogram score to evaluate patients who diagnosed with COVID-19 developed from mild or moderate to severe probability.

Methods

Patients recruitment

A single-center, retrospective, observational study was conducted from the Wuchang mobile cabin hospital built by converting public venues. We retrospectively analyzed 483 patients diagnosed with mild to moderate COVID-19 infection from Feb 6, 2020, to Feb 15, 2020. All patients diagnosed with COVID-19 infection according to the “New Coronavirus Infected Pneumonia Diagnosis and Treatment Plan (Trial Version 5)” during the study period[4]. This study was approved by the Ethics Commission of Hubei Cancer Hospital (LLHBCH2020LW-001). This was a secondary analysis. Previous research results have been announced by the Chinese Clinical Trial Registry. The clinical trial registration (http://www.chictr.org.cn) number was ChiCTR2000030858.

Data collection

Data collection such as demographic and epidemiological information, laboratory findings, nursing records, imaging information, and treatment were obtained from electronic medical records. To make sure that all patients were followed-up, two researchers independently reviewed and verified the data by telephone communication. By March 9, 2020, 62 patients who were transferred to a designated hospital.

Statistical analysis

Continuous variables were described as the mean (standard deviation [SD]). Two-tailed *t* test or Mann-Whitney test was used for continuous variables analysis. Categorical variables were compared using the $\chi 2$ test or Fisher exact test. Stepwise multivariable logistic regression was performed to identify deterioration factors associated with serious illness. A nomogram was conducted based on the results of multivariate logistic regression analysis herein. Patients were randomly divided into the training and validation samples. The predictive performance of the nomogram was evaluated via concordance index and calibration curve. To accurately decrease the overfit bias, 1000 bootstrap samples was performed. P-value < 0.05 was considered to be statistically significant. All analyses were conducted using the R statistical package (v.3.5.2; R Foundation for Statistical Computing, Vienna, Austria; https://www.r-project.org).

Results

Demographics and clinical characteristics

A total of 483 patients diagnosed mild to moderate from the mobile cabin hospital in Wuchang, Wuhan were included in this study. Among them, 62 patients developed severe or extremely critical illness during isolation observation and treatment. According to the principle of random grouping, 70% (339) of the patients are used to train the model, and 30% (144) of the patients are used to verify the model. The patient information of the two sets has been displayed through Table 1.

Table 1
Demographics and clinical characteristics in the training and validation cohort

variable	Training cohort		Validation cohort		P value
	Mild	**Severe**	**Mild**	**Severe**	
	(n = 293)	**(n = 46)**	**(n = 128)**	**(n = 16)**	
Age, y					0.94
Mean(± SD)	48.3(± 12.2)	49.0(± 13.0)	48.8(± 12.8)	45.7(± 12.7)	
Gender					0.92
female	156 (53.2%)	30 (65.2%)	71 (55.5%)	8 (50.0%)	
male	137 (46.8%)	16 (34.8%)	57 (44.5%)	8 (50.0%)	
BMI, kg/m^2					0.34
Mean(± SD)	23.3(± 3.16)	23.0(± 3.16)	23.4(± 3.13)	24.7(± 2.86)	
Exercise					0.46
No	108 (36.9%)	21 (45.7%)	55 (43.0%)	5 (31.2%)	
Yes	185 (63.1%)	25 (54.3%)	73 (57.0%)	11 (68.8%)	
Smoking					0.85
no	240 (81.9%)	31 (67.4%)	101 (78.9%)	13 (81.2%)	
yes	53 (18.1%)	15 (32.6%)	27 (21.1%)	3 (18.8%)	
Surgery					0.05
No	61 (20.8%)	2 (4.3%)	38 (29.7%)	0 (0%)	
Yes	232 (79.2%)	44 (95.7%)	90 (70.3%)	16 (100%)	
Antiviral treatment					0.39
No	97 (33.1%)	14 (30.4%)	47 (36.7%)	6 (37.5%)	
Yes	196 (66.9%)	32 (69.6%)	81 (63.3%)	10 (62.5%)	
Antibiotic treatment					0.37
No	114 (38.9%)	17 (37.0%)	56 (43.8%)	6 (37.5%)	

Abbreviations: BMI: Body Mass Index; SD: standard deviation. Hypertension: ≥140 / 90 mmHg. Normal reference value: (1) Leukocyte: Adult: (4.0–10.0) × 10 ^ 9 / L; Child: (5.0–12.0) × 10 ^ 9 / L; (2) Lymphocyte percentage (Lymph%) 20–40%; Lymphocyte absolute value (Lymph #) 1.1–3.2 × 10 ^ 9; (3) Fasting whole blood glucose: 3.9 ~ 6.1 mmol / L, 1 hour after meal: 6.7 ~ 9.4 mmol/L, 2 hours after meal: ≤7.8 mmol/L. Heart function: Tachycardia (100 beats / min). Liver function: ALT 0–46 U/L; AST 0–46 U/L. Urine infection: Creatinine (30–110 μmol/L).

Yes	179 (61.1%)	29 (63.0%)	72 (56.2%)	10 (62.5%)	
Hypertension					0.63
No	260 (88.7%)	44 (95.7%)	111 (86.7%)	16 (100%)	
Yes	33 (11.3%)	2 (4.3%)	17 (13.3%)	0 (0%)	
Diabetes					0.60
No	278 (94.9%)	43 (93.5%)	123 (96.1%)	15 (93.8%)	
Yes	15 (5.1%)	3 (6.5%)	5 (3.9%)	1 (6.2%)	
Hyperlipidemia					0.25
No	286 (97.6%)	42 (91.3%)	126 (98.4%)	16 (100%)	
Yes	7 (2.4%)	4 (8.7%)	2 (1.6%)	0 (0%)	
Malignant					0.62
No	290 (99.0%)	46 (100%)	126 (98.4%)	16 (100%)	
Yes	3 (1.0%)	0 (0%)	2 (1.6%)	0 (0%)	
Temperature,℃					1.00
Mean(± SD)	37.7(± 0.832)	37.8(± 0.918)	37.6(± 0.907)	38.3(± 0.690)	
Cough					0.21
No	117 (39.9%)	20 (43.5%)	58 (45.3%)	9 (56.2%)	
Yes	176 (60.1%)	26 (56.5%)	70 (54.7%)	7 (43.8%)	
Headache					0.19
No	228 (77.8%)	34 (73.9%)	105 (82.0%)	14 (87.5%)	
Yes	65 (22.2%)	12 (26.1%)	23 (18.0%)	2 (12.5%)	
Chest tightness					0.65
No	216 (73.7%)	29 (63.0%)	98 (76.6%)	9 (56.2%)	
Yes	77 (26.3%)	17 (37.0%)	30 (23.4%)	7 (43.8%)	
Vomiting					0.94

Abbreviations: BMI: Body Mass Index; SD: standard deviation. Hypertension: ≥140 / 90 mmHg. Normal reference value: (1) Leukocyte: Adult: (4.0–10.0) × 10 ^ 9 / L; Child: (5.0–12.0) × 10 ^ 9 / L; (2) Lymphocyte percentage (Lymph%) 20–40%; Lymphocyte absolute value (Lymph #) 1.1–3.2 × 10 ^ 9; (3) Fasting whole blood glucose: 3.9 ~ 6.1 mmol / L, 1 hour after meal: 6.7 ~ 9.4 mmol/L, 2 hours after meal: ≤7.8 mmol/L. Heart function: Tachycardia (100 beats / min). Liver function: ALT 0–46 U/L; AST 0–46 U/L. Urine infection: Creatinine (30–110 μmol/L).

No	246 (84.0%)	45 (97.8%)	109 (85.2%)	15 (93.8%)	
Yes	47 (16.0%)	1 (2.2%)	19 (14.8%)	1 (6.2%)	
Diarrhea					0.92
No	275 (93.9%)	39 (84.8%)	120 (93.8%)	13 (81.2%)	
Yes	18 (6.1%)	7 (15.2%)	8 (6.2%)	3 (18.8%)	
Leukocyte					0.18
abnormal	72 (24.6%)	18 (39.1%)	25 (19.5%)	5 (31.2%)	
normal	221 (75.4%)	28 (60.9%)	103 (80.5%)	11 (68.8%)	
Lymphocyte					0.31
abnormal	49 (16.7%)	16 (34.8%)	15 (11.7%)	7 (43.8%)	
normal	244 (83.3%)	30 (65.2%)	113 (88.3%)	9 (56.2%)	
Thrombocyte					0.45
abnormal	48 (16.4%)	11 (23.9%)	19 (14.8%)	2 (12.5%)	
normal	245 (83.6%)	35 (76.1%)	109 (85.2%)	14 (87.5%)	
Renal function					0.85
No	293 (100%)	38 (82.6%)	127 (99.2%)	14 (87.5%)	
Yes	0 (0%)	8 (17.4%)	1 (0.8%)	2 (12.5%)	
Heart function					0.56
No	289 (98.6%)	37 (80.4%)	127 (99.2%)	13 (81.2%)	
Yes	4 (1.4%)	9 (19.6%)	1 (0.8%)	3 (18.8%)	
Urine infection					0.66
No	266 (90.8%)	38 (82.6%)	117 (91.4%)	14 (87.5%)	
Yes	27 (9.2%)	8 (17.4%)	11 (8.6%)	2 (12.5%)	
Imaging test					0.27
normal	288 (98.3%)	31 (67.4%)	127 (99.2%)	12 (75.0%)	

Abbreviations: BMI: Body Mass Index; SD: standard deviation. Hypertension: ≥140 / 90 mmHg. Normal reference value: (1) Leukocyte: Adult: (4.0–10.0) × 10 ^ 9 / L; Child: (5.0–12.0) × 10 ^ 9 / L; (2) Lymphocyte percentage (Lymph%) 20–40%; Lymphocyte absolute value (Lymph #) 1.1–3.2 × 10 ^ 9; (3) Fasting whole blood glucose: 3.9 ~ 6.1 mmol / L, 1 hour after meal: 6.7 ~ 9.4 mmol/L, 2 hours after meal: ≤7.8 mmol/L. Heart function: Tachycardia (100 beats / min). Liver function: ALT 0–46 U/L; AST 0–46 U/L. Urine infection: Creatinine (30–110 μmol/L).

abnormal	5 (1.7%)	15 (32.6%)	1 (0.8%)	4 (25.0%)	
Mentality					0.92
nervous	92 (31.4%)	24 (52.2%)	46 (35.9%)	4 (25.0%)	
without nervous	201 (68.6%)	22 (47.8%)	82 (64.1%)	12 (75.0%)	

Abbreviations: BMI: Body Mass Index; SD: standard deviation. Hypertension: ≥140 / 90 mmHg. Normal reference value: (1) Leukocyte: Adult: (4.0–10.0) × 10 ^ 9 / L; Child: (5.0–12.0) × 10 ^ 9 / L; (2) Lymphocyte percentage (Lymph%) 20–40%; Lymphocyte absolute value (Lymph #) 1.1–3.2 × 10 ^ 9; (3) Fasting whole blood glucose: 3.9 ~ 6.1 mmol / L, 1 hour after meal: 6.7 ~ 9.4 mmol/L, 2 hours after meal: ≤7.8 mmol/L. Heart function: Tachycardia (100 beats / min). Liver function: ALT 0–46 U/L; AST 0–46 U/L. Urine infection: Creatinine (30–110 μmol/L).

Multivariate logistic regression analysis of deterioration factors

In order to establish a reliable prediction model, a total of thirty variables were included in this study for screening. After multivariate logistics stepwise regression analysis, there were seven variables (hyperlipidemia, vomiting, diarrhea, lymphocyte, imaging, mentality) that met the modeling criteria. Considering that age has a greater influence on the robustness of the model, for which also included in this model construction, and the relevant variables have been shown in Table 2.

Table 2
Multivariable logistic regression analysis for deteriorating factors for predicting development of severe cases

variables†	Odds ratio	95%CI	P value
Age	0.58	0.30–1.14	< 0.10£
Hyperlipidemia (yes vs no)	8.49	1.40-51.37	< 0.05
Vomiting (yes vs no)	0.003	0.00-0.07	< 0.01
Diarrhea (yes vs no)	3.51	1.17–10.58	< 0.05
Lymphocyte (normal vs abnormal)	2.65	1.13–6.18	< 0.05
Imaging (normal vs abnormal)	220.53	25.94-1875.10	< 0.01
Mentality (normal vs nervous)	3.27	1.45–7.34	< 0.01
†. adjusted model:			
Model1, body temperature was adjusted. C-index:0.874.			
Model2, gender was adjusted. C-index:0.873.			
Model3, age was adjusted. C-index:0.858.			
£. age is considered as significant variable for great influence on model adjustment.			

Construction of a nomogram score based on risk factors

The deterioration trajectories associated with serious illness was constructed based on the coefficients from the seven meaningful variables. According to the construction principle of nomogram score, each variable will be given different scores and weights. Patients can evaluate the score of each variable in turn according to their clinical characteristics and auxiliary examination results, and then summarize according to the total score of seven variables. Based on the total score, the patient can determine the probability of developing severe or extremely critical illness. The nomogram score was shown in Fig. 1.

Model accuracy evaluation by calibration curve and ROC analysis

In order to further verify the robustness of the model, the calibration curve and ROC curve are used to evaluate the pros and cons of the model. The results show that the prediction capabilities of the training set and the test set are highly consistent. The ROC curve shows that the area under the curve of the training set and the validation set are 0.873 and 0.813, respectively. The detailed information has been shown in Fig. 2 and Fig. 3.

Discussion

The use of the nomogram score in estimating the risk of patients diagnosed with mild or moderate harboring severe probability to direct clinical treatment is a new concept[7]. Previous studies have attempted to use multivariate risk analysis to evaluate patient's condition changes[8–10]. However, the significance of accurately assessing that a patient may develop into severe or extremely critical illness is still lacking.

In the deteriorating trajectory prediction nomogram, the patient's subjective feelings and auxiliary diagnosis results were included in the model. These variables are versatile and quantifiable, so that they can be directly used for patient self-evaluation. Previous study has reported that age could act as an important independent predictor of mortality in SARS and MERS[11, 12]. However, our study showed that age was not associated with severity development instead of contributing to the prediction model, which demonstrated that age should not be neglected in the disease progression. With the help of the clinical prediction model we established, patients can predict the probability of their own disease progression in advance, and also optimize the allocation of medical resources, which helps clinicians to optimize and make effective treatment decisions.

Similarly, our research also had several limitations as follows. First, this was a retrospective cohort study. Although based on the real world, the clinical information of patients is inevitably biased. Second, we mainly used single-center data to build up a prediction model, and also verify the robustness of the model internally, but external data is still required for model verification, especially multi-center clinical prediction is necessary. Third, this study mainly focused on the patient's clinical information and auxiliary

examination, and the patient's subsequent treatment effects are also needed to be fully evaluated in order to obtain a better predictive effect.

Conclusion

In this study, we established a clinical prediction model that can predict the development of mild to moderate patients to severe or extremely critical. Evaluating the deterioration trajectory of COVID-19 patients can help optimize the allocation of scarce medical resources.

Abbreviations

SD: standard deviation; ROC curve: receiver operating characteristic curve; ChiCTR: Chinese Clinical Trial Registry.

Declarations

Ethics approval and consent to participate

This study was approved by the Ethics Commission of Hubei Cancer Hospital (LLHBCH2020LW-001).

Consent for publication

This study was approved for publication by the Chinese Clinical Trial Registry (ChiCTR2000030858).

Availability of data and materials

The data connection:
tiejun, Wang (2020), COVID-19, Dryad, Dataset, https://doi.org/10.5061/dryad.kwh70rz13

Competing interests

We declare no competing interests.

Funding

This study was supported by grants from the National Natural Foundation of China (No.81172467).

Author contributions

Wei Chen, Menglin Zhu and Tiejun Wang developed the original concept and study design. Jian Li, Cuiping Pan, Demian Zhao, Yuting Jin generated the initial draft of the manuscript. Manxiu Li, Shun Wu and Yaojun Feng completed the data analysis. All authors contributed to the interpretation of the findings and approved the final manuscript version.

Acknowledgements

Not applicable

References

1. Kannan, S., et al., *COVID-19 (Novel Coronavirus 2019) - recent trends.* Eur Rev Med Pharmacol Sci, 2020. **24**(4): p. 2006-2011.
2. Ahn, D.G., et al., *Current Status of Epidemiology, Diagnosis, Therapeutics, and Vaccines for Novel Coronavirus Disease 2019 (COVID-19).* J Microbiol Biotechnol, 2020. **30**(3): p. 313-324.
3. Zhou, F., 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): p. 1054-1062.
4. Chen, S., et al., *Fangcang shelter hospitals: a novel concept for responding to public health emergencies.* Lancet, 2020. **395**(10232): p. 1305-1314.
5. Li, Q., et al., *Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus-Infected Pneumonia.* N Engl J Med, 2020. **382**(13): p. 1199-1207.
6. Liang, W., 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.
7. Iasonos, A., et al., *How to build and interpret a nomogram for cancer prognosis.* J Clin Oncol, 2008. **26**(8): p. 1364-70.
8. Ali, H., et al., *Survival rate in acute kidney injury superimposed COVID-19 patients: a systematic review and meta-analysis.* Ren Fail, 2020. **42**(1): p. 393-397.
9. Alqahtani, J.S., et al., *Prevalence, Severity and Mortality associated with COPD and Smoking in patients with COVID-19: A Rapid Systematic Review and Meta-Analysis.* PLoS One, 2020. **15**(5): p. e0233147.
10. Lippi, G., J. Wong, and B.M. Henry, *Hypertension in patients with coronavirus disease 2019 (COVID-19): a pooled analysis.* Pol Arch Intern Med, 2020. **130**(4): p. 304-309.
11. Choi, K.W., et al., *Outcomes and prognostic factors in 267 patients with severe acute respiratory syndrome in Hong Kong.* Ann Intern Med, 2003. **139**(9): p. 715-23.
12. Matsuyama, R., et al., *Clinical determinants of the severity of Middle East respiratory syndrome (MERS): a systematic review and meta-analysis.* BMC Public Health, 2016. **16**(1): p. 1203.

Figures

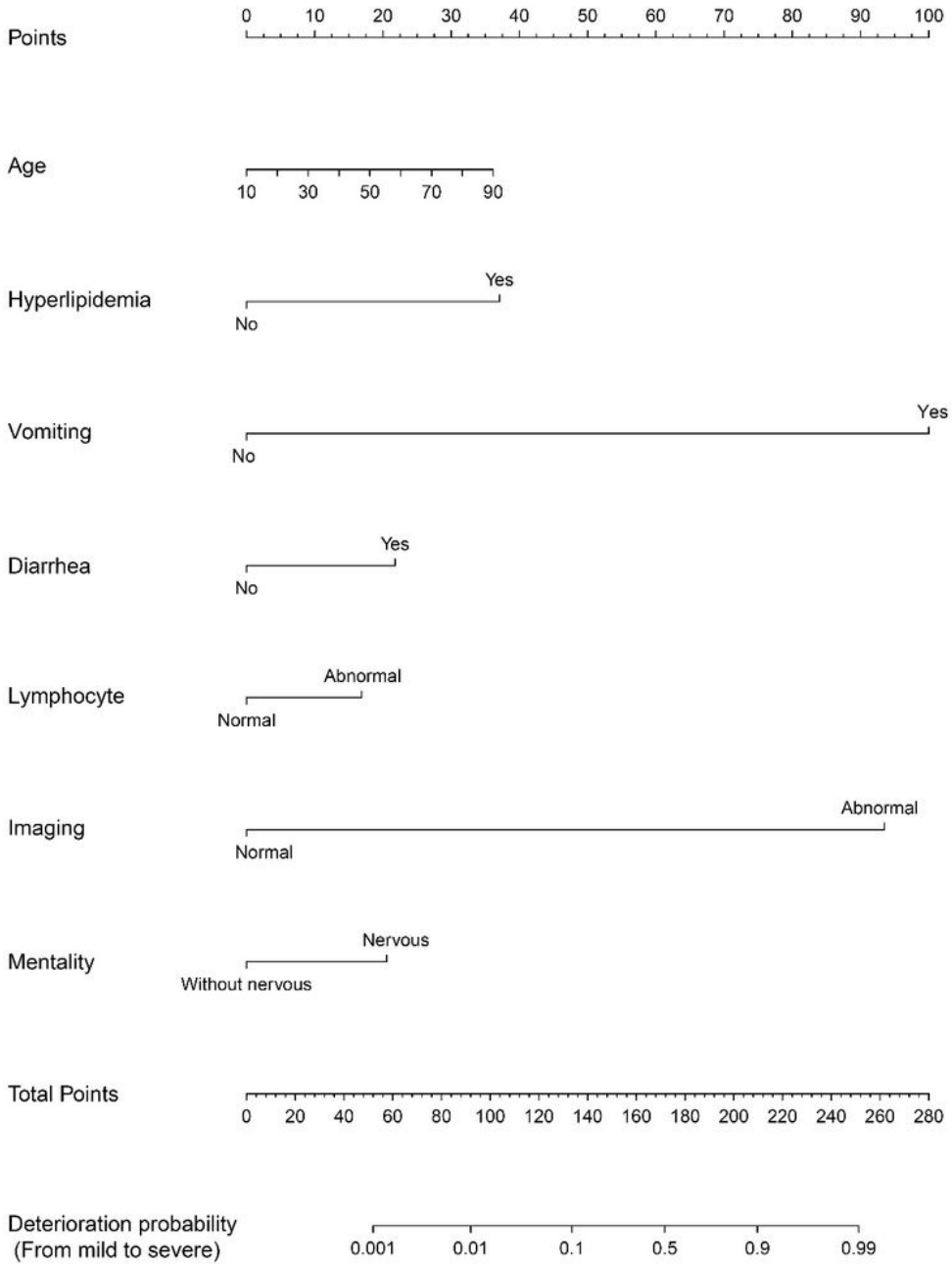

Figure 1

Nomogram score for predicting development of severe probability

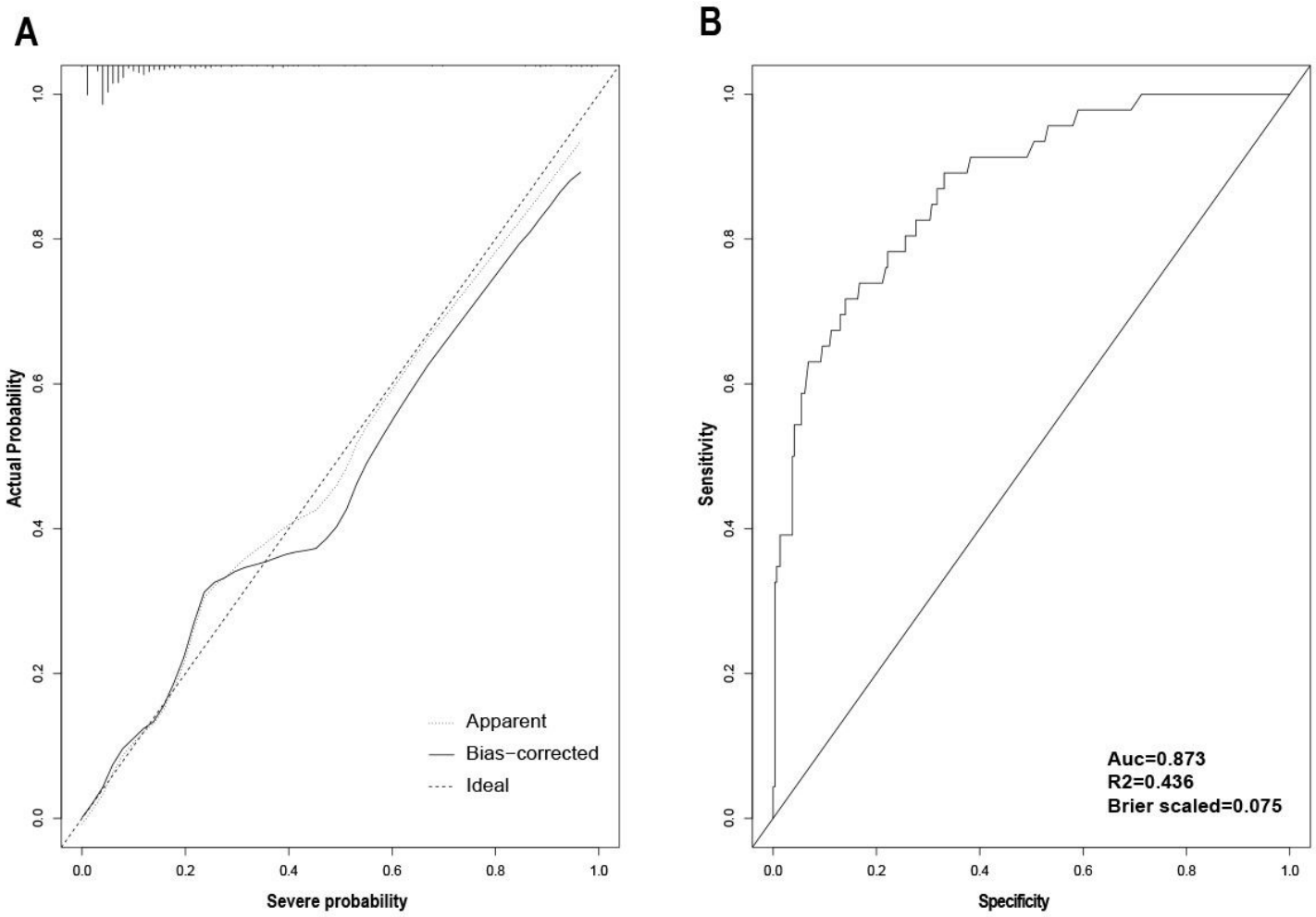

Figure 2

Consistency and sensitivity analysis of nomogram score in the training cohort A. The prediction accuracy of continuous correction curve detection model. B. Consistency and sensitivity of ROC curve detection model.

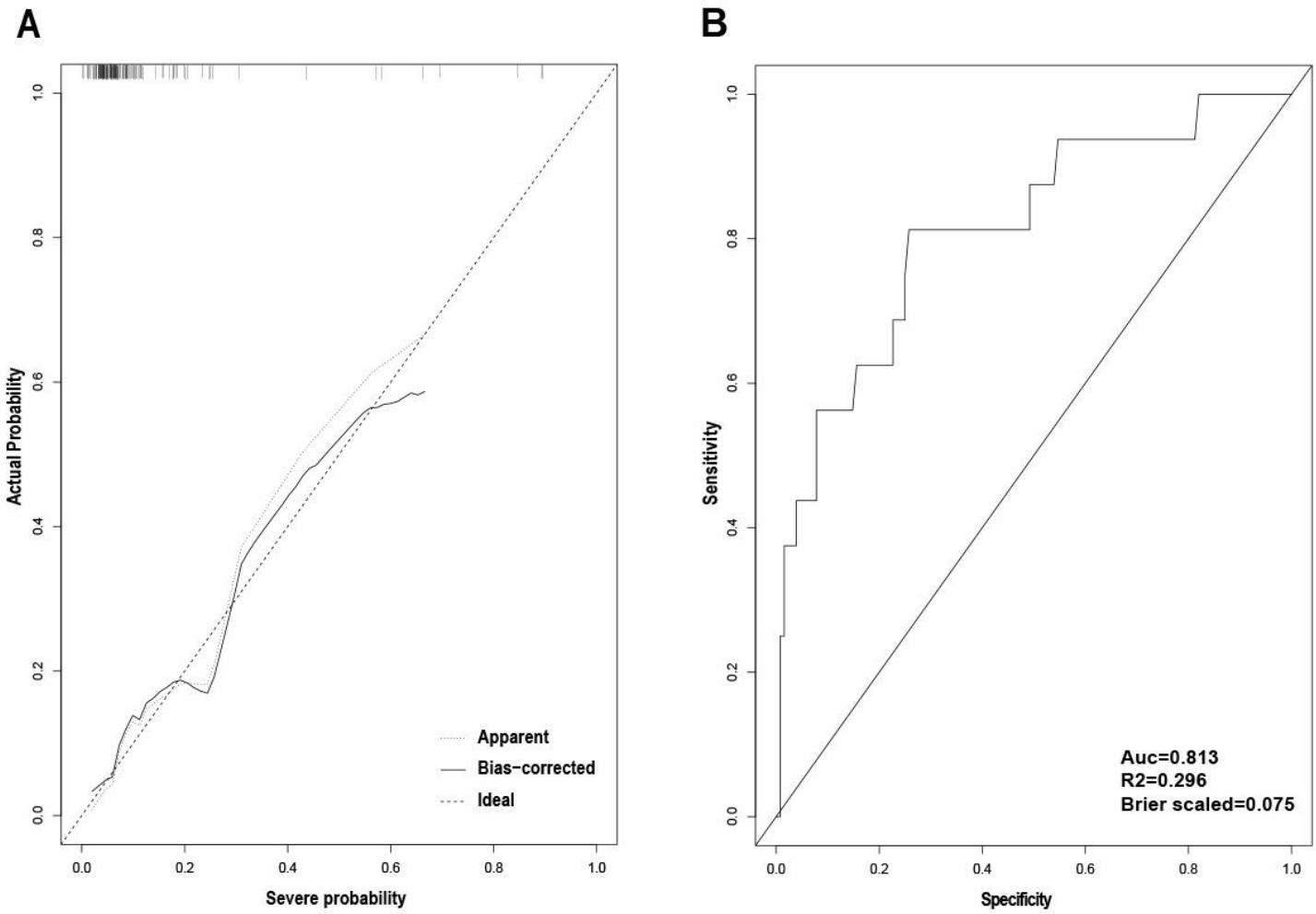

Figure 3

Consistency and sensitivity analysis of nomogram score in the validation cohort. A. The prediction accuracy of continuous correction curve detection model. B. Consistency and sensitivity of ROC curve detection model.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- highlights.docx
- titlepage.docx