Engineering 8 (2022) 122–129

Contents lists available at ScienceDirect

Engineering

journal homepage: www.elsevier.com/locate/eng

Research
Coronavirus Disease 2019—Article

A Predictive Nomogram for Predicting Improved Clinical Outcome Probability in Patients with COVID-19 in Zhejiang Province, China

Jiaojiao Xie [a,#], Ding Shi [a,#], Mingyang Bao [b,#], Xiaoyi Hu [c,#], Wenrui Wu [a], Jifang Sheng [a], Kaijin Xu [a], Qing Wang [a], Jingjing Wu [a], Kaicen Wang [a], Daiqiong Fang [d], Yating Li [a], Lanjuan Li [a,*]

[a] State Key Laboratory for Diagnosis and Treatment of Infectious Diseases, National Clinical Research Center for Infectious Diseases, Collaborative Innovation Center for Diagnosis and Treatment of Infectious Diseases, The First Affiliated Hospital, College of Medicine, Zhejiang University, Hangzhou 310003, China
[b] State Key Laboratory of Genetic Engineering, Institute of Biostatistics, School of Life Sciences, Fudan University, Shanghai 200433, China
[c] Division of Hepatobiliary and Pancreatic Surgery, Department of Surgery, Key Lab of Combined Multi-organ Transplantation of the Ministry of Health, The First Affiliated Hospital, College of Medicine, Zhejiang University, Hangzhou 310003, China
[d] Division of of Endocrinology and Metabolism, Department of Internal Medicine System, The First Affiliated Hospital, College of Medicine, Zhejiang University, Hangzhou 310003, China

ARTICLE INFO

Article history:
Received 19 March 2020
Revised 14 April 2020
Accepted 3 May 2020
Available online 6 June 2020

Keywords:
Coronavirus disease 2019 (COVID-19)
Nomogram
Patient-relevant outcome

ABSTRACT

The aim of this research was to develop a quantitative method for clinicians to predict the probability of improved prognosis in patients with coronavirus disease 2019 (COVID-19). Data on 104 patients admitted to hospital with laboratory-confirmed COVID-19 infection from 10 January 2020 to 26 February 2020 were collected. Clinical information and laboratory findings were collected and compared between the outcomes of improved patients and non-improved patients. The least absolute shrinkage and selection operator (LASSO) logistics regression model and two-way stepwise strategy in the multivariate logistics regression model were used to select prognostic factors for predicting clinical outcomes in COVID-19 patients. The concordance index (C-index) was used to assess the discrimination of the model, and internal validation was performed through bootstrap resampling. A novel predictive nomogram was constructed by incorporating these features. Of the 104 patients included in the study (median age 55 years), 75 (72.1%) had improved short-term outcomes, while 29 (27.9%) showed no signs of improvement. There were numerous differences in clinical characteristics and laboratory findings between patients with improved outcomes and patients without improved outcomes. After a multi-step screening process, prognostic factors were selected and incorporated into the nomogram construction, including immunoglobulin A (IgA), C-reactive protein (CRP), creatine kinase (CK), acute physiology and chronic health evaluation II (APACHE II), and interaction between CK and APACHE II. The C-index of our model was 0.962 (95% confidence interval (CI), 0.931–0.993) and still reached a high value of 0.948 through bootstrapping validation. A predictive nomogram we further established showed close performance compared with the ideal model on the calibration plot and was clinically practical according to the decision curve and clinical impact curve. The nomogram we constructed is useful for clinicians to predict improved clinical outcome probability for each COVID-19 patient, which may facilitate personalized counselling and treatment.

© 2022 THE AUTHORS. Published by Elsevier LTD on behalf of Chinese Academy of Engineering and Higher Education Press Limited Company. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

Coronavirus disease 2019 (COVID-19), caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), has been regarded as a pandemic disaster, while cases continue to rise globally [1–5]. SARS-CoV-2, which has been identified as a β-coronavirus, a clade in the rotavirus subgenus belonging to subfamily of the orthodox coronavirus, has a similar phylogeny to two other β-coronavirus: severe acute respiratory syndrome coronavirus (SARS-CoV) and Middle East respiratory syndrome coronavirus (MERS-CoV). β-coronaviruses that are zoonotic in origin have been linked to potentially fatal illness during outbreaks in 2003 and 2012, respectively [6,7].

* Corresponding author.
E-mail address: ljli@zju.edu.cn (L. Li).
# These authors contributed equally to this work.

https://doi.org/10.1016/j.eng.2020.05.014
2095-8099/© 2022 THE AUTHORS. Published by Elsevier LTD on behalf of Chinese Academy of Engineering and Higher Education Press Limited Company.
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

By 3 March 2020, the total number of infections and deaths caused by COVID-19 had risen sharply to over 83 000 worldwide. The incubation period is 2–14 d from COVID-19 infection to the onset of symptoms. Clinical manifestations are very similar to those of SARS, including fever, cough, nausea, and vomiting. However, some COVID-19 patients have no fever or radiological abnormalities at the beginning, which complicates the diagnosis. Laboratory findings indicate that lymphocytopenia (83.2%), thrombocytopenia (36.2%), leukopenia (33.7%), and elevated levels of C-reactive protein (CRP) are the most common characteristics among patients with COVID-19 [8]. According to current evidence, the mortality rate of COVID-19 is about 3%, and deaths mainly occur in older patients or those with coexisting diseases [9,10]. There is no doubt that early management may make a significant contribution to reducing mortality. Previous studies have demonstrated the general epidemiological and clinical characteristics as well as potential rapid diagnostics, vaccine, and therapeutics of COVID-19 pneumonia [11,12]. However, the association of demographic traits, laboratory indicator levels, and examination results with outcome improvement remains unclear. Furthermore, current exploration of the underlying factors for early intervention and prognosis of COVID-19 is still insufficient.

In our study, we aimed to develop a quantitative method for clinicians to predict the probability of improved prognosis in each COVID-19 patient. A total of 104 patients with laboratory-confirmed COVID-19 infection in Zhejiang Province were divided into two groups based on whether the outcome improved. A least absolute shrinkage and selection operator (LASSO) logistics regression model was used to select the optimal prognostic indicators from the clinical characteristics and laboratory findings of COVID-19 cases. A further filter was conducted by a two-way stepwise strategy in the multivariate logistics regression model. The final COVID-19-related predictive model consists of five prognostic factors. A nomogram was eventually constructed to predict the probability of outcome improvement by incorporating these variables. This work has created an effective nomogram for improved prediction of COVID-19 patients, which can be used to optimize treatment strategy.

2. Methods

2.1. Data sources

We obtained the medical records and compiled data for patients with laboratory-confirmed COVID-19 according to World Health Organization (WHO) interim guidance [13] from 10 January 2020 to 26 February 2020. All of the cases enrolled in this study were confirmed by the laboratory to be COVID-19 infection based on real-time reverse transcriptase-polymerase chain reaction (RT-PCR) assay of nasal and pharyngeal swab specimens. We collected data on 104 patients admitted to hospital with laboratory-confirmed COVID-19 infection at the First Affiliated Hospital, Zhejiang University, Zhejiang Province, China. Information was collected on electronic medical records, interviews of investigators, and hospital admissions. The data were reviewed by a trained team of physicians. The days of onset of symptom to diagnosis were counted from the illness onset to the laboratory confirmation of COVID-19 infection. We defined the degree of severity of COVID-19 at the time of admission using the acute physiology and chronic health evaluation II (APACHE II) [14]. Exposure history means having close contact (gathering, living, or working together) with individuals with confirmed or suspected COVID-19 infection during the two weeks before illness onset. Familial clusters were defined as patients who infected others in their families. All patients were classified into three grades (moderate/severe/critical) based on the *Guidelines for diagnosis and treatment of novel coronavirus pneumonia (trail version 7)* [15]. Fever was defined as an axillary temperature of 37.3 °C or higher. Chest radiography or computed tomography (CT) and all laboratory testing were performed according to the clinical care needs for the patient. We determined the presence of a radiologic abnormality on the findings of bilateral or multiple lobular or subsegmental areas of consolidation or bilateral ground glass, and ranked the scores based on the numbers of involved pulmonary segments: 1 (normal); 2 (1–2); 3 (3–5); 4 (> 5). All measures of arterial pressure and partial pressure of carbon dioxide were recorded by professional physicians.

2.2. Laboratory confirmation

Sputum and throat swab specimens were collected from all patients at admission. Laboratory confirmation of the virus was performed by RT-PCR assay for COVID-19 ribonucleic acid (RNA) within 3 h. Virus detection was repeated twice every 24 h. All laboratory tests were performed according to the clinical care needs of the patient. Laboratory assessments consisted of a complete blood count; blood chemical analysis; coagulation testing; assessment of liver and renal function; and measures of CRP, procalcitonin (PCT), lactate dehydrogenase, creatine kinase (CK), inflammatory cytokines, complement, and immunoglobulin.

2.3. Statistical analysis

Patients were divided into two groups based on whether the outcomes improved. The following two conditions were defined as outcome improvement: ① Severe patients who were admitted to the intensive care unit (ICU) at the beginning of hospital admission alleviated after treatment and were transferred out of the ICU to general isolated wards; and ② those with mild illness at the time of hospital admission were discharged or were going to be discharged at the end of the follow-up. Conversely, patients who received continuous treatment in the ICU or subsequent transition to the ICU due to exacerbation were considered to have non-improved outcomes.

Normally distributed continuous variables were described as means with standard deviations (SD), and parametric *t*-tests were used to test for statistical significance between the two groups; otherwise, medians with interquartile range (IQR) and non-parametric Mann–Whitney *U* tests were applied for variable description and two comparisons, respectively. For categorical variables, we expressed the numbers and percentages of patients in each category. Proportions were compared using the χ^2 test, with Yates' correction or Fisher's exact test.

LASSO logistics regression analysis was performed to select the optimal prognostic indicators from demographic characteristics, examinations, coexisting conditions, symptoms, and laboratory findings for COVID-19 patients. The logistics regression model with the LASSO penalty successfully achieved dimensionality reduction. The optimal value of the penalty parameter λ was adopted and variables with nonzero coefficients in the model were selected. A further filter was conducted by a two-way stepwise strategy in the multivariate logistics regression model. Interaction between every two pair of variables was taken into account. Moreover, the concordance index (C-index) was computed to evaluate the discrimination performance of our model. A relatively corrected C-index was calculated by 1000 bootstrap resampling for validation. Given the wide range of laboratory indicators, we further divided them into quartiles as categorical variables in order to assess their association with the probability of improvement. In addition, patients were classified into four age groups: < 40, 40–54, 55–69, and $\geq$ 70 years, in order to investigate the effects of age on the outcome.

After a multi-step screening process, the final prognostic factors were used to construct a nomogram for predicting the probability of outcome improvement. According to the regression coefficient, each variable that was included corresponded to a point at each value. A total point was equal to the sum of the points of all variables for each patient. The relationship between the total points and the probability of outcome improvement was visualized on the bottom of the nomogram. Calibration curves were subsequently drawn to assess the agreement between the nomogram-predicted probability and the actual proportion. As a reference line, the diagonal represents the best prediction. Moreover, we performed a decision curve analysis to determine whether our established nomogram was suitable for clinical utility by estimating the net benefits at different threshold probabilities. The clinical impact curve was drawn to predict improved probability stratification for a population size as 1000.

A two-sided P value < 0.05 was considered to be statistically significant. All statistical analyses were performed using R 3.6.1 software.

3. Results

3.1. Patients' characteristics

Clinical characteristics were collected from 104 patients with laboratory-confirmed COVID-19 who were admitted to our hospital by 26 February 2020 (Table 1). The median age was 55 years (IQR: 43–64) and 60.6% of patients were male (63). The median duration from the onset of symptoms to diagnosis was 5 d (IQR: 2–7). Of the 104 patients, 80 (76.9%) had been exposed to individuals with confirmed COVID-19 infection. Half of the cases showed a familial cluster. After a preliminary medical examination, we detected intestinal flora disorders, bacterial infection, fecal RNA positive, and acute respiratory distress syndrome (ARDS) in 9 (8.7%), 13 (12.5%), 29 (27.9%), and 16 (15.4%) patients, respectively. The median APACHE II score was 6 (IQR: 3–11) on the day of hospital admission, and more than half of the patients were assessed as grade 4 from the results of the chest CT scan. Moderate, severe, and critical patients each accounted for approximately one third of

Table 1
Demographic and clinical characteristics of 104 patients with laboratory-confirmed COVID-19.

Characteristic	Outcome			P value
	All patients (n = 104)	Improvement (n = 75)	Without improvement (n = 29)	
APACHE II, median (IQR)	6 (3–11)	5 (2.5–7)	12 (11–15)	< 0.001
Onset to diagnosis (d), median (IQR)	5 (2–7)	5 (2.5–8)	5 (1–7)	0.512
Age (years), median (IQR)	55 (43–64)	51 (38–59)	66 (59–80)	< 0.001
Gender				> 0.999
Female	41 (39.4%)	30 (40.0%)	11 (37.9%)	
Male	63 (60.6%)	45 (60.0%)	18 (62.1%)	
Body mass index				0.580
< 18	3 (2.9%)	2 (2.7%)	1 (3.4%)	
18.0–23.9	38 (36.5%)	28 (37.3%)	10 (34.5%)	
24.0–27.9	43 (41.3%)	35 (46.7%)	8 (27.6%)	
≥28	12 (11.5%)	8 (10.7%)	4 (13.8%)	
Exposure	80 (76.9%)	56 (74.7%)	24 (82.8%)	0.536
Familial cluster	52 (50.0%)	39 (52.0%)	13 (44.8%)	0.662
Grade				< 0.001
Moderate	37 (35.6%)	36 (48.0%)	1 (3.4%)	
Severe	35 (33.7%)	32 (42.7%)	3 (10.3%)	
Critical	32 (30.8%)	7 (9.3%)	25 (86.2%)	
Examinations				
Intestinal flora disorders	9 (8.7%)	6 (8.0%)	3 (10.3%)	0.707
Bacterial infection	13 (12.5%)	5 (6.7%)	8 (27.6%)	0.007
Fecal RNA positive	29 (27.9%)	19 (25.3%)	10 (34.5%)	0.491
ARDS	16 (15.4%)	4 (5.3%)	12 (41.4%)	< 0.001
Classification of chest CT scan				0.009
1	11 (10.6%)	11 (14.7%)	0 (0%)	
2	20 (19.2%)	18 (24.0%)	2 (6.9%)	
3	19 (18.3%)	13 (17.3%)	6 (20.7%)	
4	54 (51.9%)	33 (44.0%)	21 (72.4%)	
Coexisting conditions				
Smoking	14 (13.5%)	9 (12.0%)	5 (17.2%)	0.527
Alcohol	8 (7.7%)	7 (9.3%)	1 (3.4%)	0.438
Hypertension	39 (37.5%)	21 (28.0%)	18 (62.1%)	0.003
Diabetes	17 (16.3%)	9 (12.0%)	8 (27.6%)	0.075
Fatty liver	16 (15.4%)	9 (12.0%)	7 (24.1%)	0.139
Other	31 (29.8%)	19 (25.3%)	12 (41.4%)	0.172
Symptoms				
Fever	88 (84.6%)	65 (86.7%)	23 (79.3%)	0.373
Nausea	6 (5.8%)	5 (6.7%)	1 (3.4%)	> 0.999
Emesis	3 (2.9%)	2 (2.7%)	1 (3.4%)	> 0.999
Headache	8 (7.7%)	8 (10.7%)	0 (0%)	0.102
Cough	84 (80.8%)	60 (80.0%)	24 (82.8%)	> 0.999
Expectoration	49 (47.1%)	31 (41.3%)	18 (62.1%)	0.093
Chest distress	47 (45.2%)	30 (40.0%)	17 (58.6%)	0.136
Abdominal pain	2 (1.9%)	2 (2.7%)	0 (0%)	> 0.999
Diarrhea	13 (12.5%)	7 (9.3%)	6 (20.7%)	0.182
Myalgia	22 (21.2%)	17 (22.7%)	5 (17.2%)	0.605
Pharyngalgia	5 (4.8%)	3 (4.0%)	2 (6.9%)	0.617
Fatigue	29 (27.9%)	17 (22.7%)	12 (41.4%)	0.096

Note: The body mass index values were missing in eight patients.
ARDS: acute respiratory distress syndrome.

the total, respectively. Furthermore, hypertension (39 (37.5%)) was the most common coexisting medical condition, and 31 (29.8%) patients suffered from other comorbidities, such as stroke, coronary heart disease, and dyslipidemia. The most common symptom at the onset of illness was fever (88 (84.6%)), followed by cough (84 (80.8%)), expectoration (49 (47.1%)), and chest distress (47 (45.2%)).

Of these patients, 75 (72.1%) had improved outcomes by 26 February 2020, while another 29 (27.9%) showed no signs of improvement. Compared with the improved patients, those with developing illness had significantly higher APACHE II scores (12 (IQR: 11–15) vs 5 (IQR: 2.5–7); $P < 0.001$) and were significantly older (66 years (IQR: 59–80) vs 51 years (IQR: 38–59); $P < 0.001$). The proportions of critical illness, bacterial infection, ARDS, and high CT classification in cases without improvement were higher than those in cases with improvement. Patients without improvement were more likely to have hypertension than those with improvement (18 (62.1%) vs 21 (28.0%); $P = 0.003$). However, no significant difference in symptoms between the two groups of patients was observed (Table 1).

3.2. Laboratory findings

There were numerous differences in the laboratory findings between the improved and non-improved patients (Table 2). The median of the ratio of partial pressure of oxygen (PaO_2) to fraction of inspired oxygen (FIO_2) was significantly higher in the improved samples than in the non-improved samples (288.8 (IQR: 234.2–390.7) vs 205.9 (IQR: 141.8–289.4); $P = 0.005$). In terms of routine blood tests, there were significant differences in hemoglobin, red

Table 2
Laboratory findings of 104 patients with confirmed COVID-19 infection.

Laboratory findings	Normal range	Outcomes			*P* value
		All patients (*n* = 104)	Improvement (*n* = 75)	Without improvement (*n* = 29)	
Normally distributed variables, mean ± SD					
Mean arterial pressure (mmHg)	70.0–105.0	96.0 ± 12.5	95.3 ± 13.4	97.8 ± 9.7	0.297
Partial pressure of carbon dioxide (mmHg)	35.0–48.0	36.3 ± 4.9	36.7 ± 4.2	35.3 ± 6.1	0.298
Red blood cell count ($\times 10^{12}$ L^{-1})	3.7–5.1	4.4 ± 0.6	4.5 ± 0.6	4.2 ± 0.6	0.030
Hemoglobin ($g \cdot L^{-1}$)	113.0–151.0	133.7 ± 17.3	135.8 ± 16.6	128.2 ± 17.7	0.044
Albumin ($g \cdot L^{-1}$)	34.0–54.0	38.3 ± 5.8	39.9 ± 5.4	34.3 ± 4.5	< 0.001
IgA ($mg \cdot dL^{-1}$)	76.0–390.0	216.3 ± 83.1	204.9 ± 79.3	245.4 ± 85.3	0.032
High density lipoprotein ($mmol \cdot L^{-1}$)	0.9–2.4	1.0 ± 0.3	1.0 ± 0.3	1.0 ± 0.2	0.834
Low density lipoprotein ($mmol \cdot L^{-1}$)	1.3–3.3	2.0 ± 0.6	2.1 ± 0.5	1.9 ± 0.6	0.140
Non-normally distributed variables, median (IQR)					
PaO_2:FIO_2	400.0–500.0	270.6 (180.8–374.7)	288.8 (234.2–390.7)	205.9 (141.8–289.4)	0.005
White blood cell count ($\times 10^9$ L^{-1})	4.0–10.0	5.9 (4.1–9.3)	5.2 (3.8–8.1)	7.9 (4.8–11.6)	0.009
Platelet count ($\times 10^9$ L^{-1})	101.0–320.0	186.5 (147.8–237.8)	193.0 (160.5–252.0)	175.0 (136.0–191.0)	0.057
Neutrophil count ($\times 10^9$ L^{-1})	2.0–7.0	4.4 (2.8–7.9)	3.6 (2.5–6.7)	7.2 (4.4–11.0)	< 0.001
Lymphocyte count ($\times 10^9$ L^{-1})	0.8–4.0	0.8 (0.5–1.2)	0.9 (0.6–1.3)	0.5 (0.4–0.8)	0.002
Lactic acid ($mmol \cdot L^{-1}$)	0.5–1.7	1.7 (1.3–2.2)	1.6 (1.3–2.1)	1.7 (1.2–2.2)	0.856
D-dimer ($mg \cdot L^{-1}$)	0–0.7	383.0 (220.5–761.0)	323.0 (196.0–621.5)	604.0 (388.0–1023.0)	0.005
International normalized ratio	0.80–1.30	0.98 (0.95–1.03)	0.97 (0.94–1.03)	1.00 (0.97–1.05)	0.039
Prothrombin time (s)	9.4–12.5	11.8 (11.4–12.3)	11.6 (11.3–12.3)	12.0 (11.6–12.3)	0.053
Activated partial thromboplastin time (s)	25.1–36.5	31.9 (29.1–35.4)	31.3 (29.0–34.6)	33.3 (29.6–37.1)	0.057
Fibrinogen level ($g \cdot L^{-1}$)	2.0–4.0	4.3 (3.9–5.4)	4.2 (3.6–5.1)	4.7 (4.1–5.5)	0.041
Globulin ($g \cdot L^{-1}$)	20.0–35.0	29.0 (26.1–33.4)	27.8 (25.8–31.4)	31.5 (29.5–36.9)	0.002
Alanine aminotransferase ($U \cdot L^{-1}$)	7.0–40.0	21.0 (15.0–31.0)	21.0 (15.0–30.5)	21.0 (16.0–31.0)	0.584
Aspartate aminotransferase ($U \cdot L^{-1}$)	13.0–35.0	22.0 (18.0–34.0)	20.0 (16.0–29.0)	29.0 (22.0–41.0)	0.003
Alkaline phosphatase ($U \cdot L^{-1}$)	35.0–100.0	62.0 (53.8–77.0)	62.0 (52.5–76.5)	60.0 (54.0–77.0)	0.643
Gamma-glutamyltransferase ($U \cdot L^{-1}$)	7.0–45.0	27.0 (18.0–52.0)	26.0 (17.5–44.0)	35.0 (21.0–72.0)	0.074
Total bilirubin ($\mu mol \cdot L^{-1}$)	0–21.0	11.2 (7.9–16.3)	11.1 (7.1–15.6)	12.1 (8.0–18.7)	0.411
Direct bilirubin ($mmol \cdot L^{-1}$)	0–8.0	4.9 (3.3–7.8)	4.7 (3.0–7.2)	5.1 (4.5–10.0)	0.033
Creatinine ($\mu mol \cdot L^{-1}$)	41.0–73.0	74.5 (62.8–89.3)	73.0 (60.0–87.5)	84.0 (68.0–103.0)	0.022
Uric acid ($\mu mol \cdot L^{-1}$)	155.0–357.0	245.0 (193.5–303.5)	240.0 (192.0–290.5)	256.0 (214.0–345.0)	0.319
Triglyceride ($mmol \cdot L^{-1}$)	0.3–1.7	1.2 (0.9–1.7)	1.2 (0.9–1.7)	1.3 (0.9–1.7)	0.943
Total cholesterol ($mmol \cdot L^{-1}$)	3.1–5.9	3.7 (3.3–4.2)	3.8 (3.3–4.2)	3.6 (3.1–4.1)	0.572
Glomerular filtration rate ($mL \cdot min^{-1}$)	NA	92.2 (73.6–103.0)	96.4 (82.3–108.4)	75.8 (55.9–92.0)	< 0.001
Glucose ($mmol \cdot L^{-1}$)	3.9–6.1	7.5 (5.4–8.9)	6.8 (5.2–8.4)	8.8 (6.7–11.1)	0.003
Potassium ($mmol \cdot L^{-1}$)	3.5–5.3	3.8 (3.5–4.2)	3.8 (3.5–4.2)	3.8 (3.5–4.0)	0.589
Sodium ($mmol \cdot L^{-1}$)	137.0–147.0	139.0 (137.0–141.0)	140.0 (137.0–141.0)	137.0 (134.0–139.0)	0.005
Lactate dehydrogenase ($U \cdot L^{-1}$)	120–250	253.5 (210.5–342.5)	229.0 (195.5–301.5)	344.0 (283.0–419.0)	< 0.001
Hydroxybutyrate dehydrogenase ($U \cdot L^{-1}$)	72.0–182.0	221.0 (174.3–276.0)	200.0 (164.5–252.0)	276.0 (236.0–331.0)	< 0.001
CK ($U \cdot L^{-1}$)	40.0–200.0	69.5 (48.8–120.5)	66.0 (48.0–99.5)	101.0 (56.0–244.0)	0.011
CKMB ($U \cdot L^{-1}$)	0–24.0	20.0 (16.0–24.0)	20.0 (15.0–23.0)	22.0 (17.0–27.0)	0.018
CRP ($mg \cdot L^{-1}$)	5.0–10.0	24.1 (9.1–50.2)	15.6 (6.9–33.2)	51.2 (29.4–93.8)	< 0.001
Troponin I ($ng \cdot mL^{-1}$)	0–0.461	0.004 (0.002–0.008)	0.003 (0.001–0.005)	0.010 (0.004–0.021)	< 0.001
PCT ($ng \cdot mL^{-1}$)	0–0.05	0.05 (0.03–0.09)	0.05 (0.03–0.07)	0.11 (0.05–0.32)	< 0.001
IL-2 ($pg \cdot mL^{-1}$)	0–4.13	0.95 (0.76–1.50)	0.95 (0.77–1.61)	0.95 (0.76–1.40)	0.859
IL-4 ($pg \cdot mL^{-1}$)	0–8.37	1.77 (1.40–1.77)	1.77 (1.40–1.77)	1.77 (1.40–1.77)	0.852
IL-6 ($pg \cdot mL^{-1}$)	0–6.6	21.9 (8.5–57.1)	16.3 (7.6–37.3)	47.2 (23.3–80.3)	0.001
IL-10 ($pg \cdot mL^{-1}$)	0–2.3	4.5 (2.9–7.8)	3.9 (2.9–7.2)	6.7 (4.8–10.1)	0.015
Tumor necrosis factor-α ($pg \cdot mL^{-1}$)	0–33.3	15.6 (7.4–54.3)	21.3 (6.9–61.9)	12.2 (9.2–26.8)	0.244
Interferon-γ ($pg \cdot mL^{-1}$)	0–20.6	9.2 (5.2–28.6)	9.9 (5.8–32.3)	9.2 (3.4–27.4)	0.332
Complement 4 ($mg \cdot dL^{-1}$)	16.0–48.0	36.0 (29.0– 42.3)	37.0 (29.0–44.0)	34.0 (30.5–41.0)	0.029
Complement 3 ($mg \cdot dL^{-1}$)	80.0–160.0	128.0 (113.8–148.3)	131.0 (117.0–150.0)	118.0 (107.5–131.5)	0.619
IgM ($mg \cdot dL^{-1}$)	40.0–345.0	78.5 (53.8–123.3)	86.0 (52.0–124.0)	72.0 (54.0–113.0)	0.257
IgG ($mg \cdot dL^{-1}$)	600–1600	1293 (1028–1758)	1243 (1000–1470)	1665 (1246–2189)	< 0.001

NA: not available; IgA: immunoglobulin A; PaO_2:FIO_2: the ratio of partial pressure of oxygen to fraction of inspired oxygen; CKMB: creatine kinase isoenzymes-myocardial band; IL: interleukin.

blood cell count, and three types of white blood cell counts. Many biochemical indicators also showed significant differences between the two groups, such as aspartate aminotransferase, creatine kinase isoenzymes-myocardial band (CKMB), glomerular filtration rate, and CK. CRP, PCT, and two inflammatory cytokines, increased significantly in cases without improvement, and were much higher than the upper limit of the normal range. Regarding immune-related proteins, more interleukin (IL)-6 and IL-10 were secreted in cases without improvement. The levels of immunoglobulin G (IgG) and immunoglobulin A (IgA) were lower in improved patients than in non-improved patients (Table 2).

3.3. Selection of prognostic predictors

All of the demographic characteristics, examinations, coexisting conditions, symptoms, and laboratory findings described above were included in the LASSO logistics regression model to screen the potential predictors. Changes in the LASSO partial likelihood deviance and coefficients with lnλ are shown in Fig. 1. As a result, 11 variables with nonzero coefficients were selected, including age, grade, headache, APACHE II, activated partial thromboplastin time, CK, CKMB, CRP, PCT, IgA, and IgG. These variables were subsequently filtered in the multivariate logistics regression model with a two-way stepwise strategy. Finally, the model including IgA, CRP, CK, and APACHE II reached the minimal Akaike information criterion (AIC), which indicated the best goodness of fit. The result of the interaction analysis revealed that there was an interaction between CK and APACHE II. Serum CK was log-transformed due to high skew to the right in this group of patients (Table 3). Furthermore, the C-index of our logistics regression model was 0.962 (95% confidence interval (CI), 0.931–0.993) and was corrected to 0.948 through bootstrapping validation, which showed that the model had good predictive power.

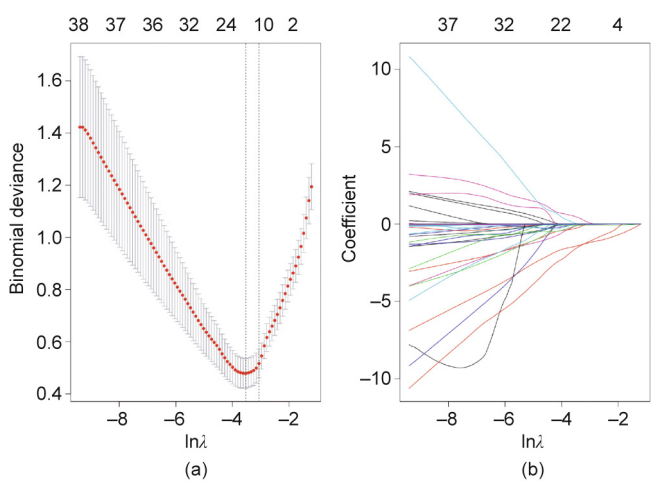

Fig. 1. LASSO logistics regression plot. (a) Plot of partial likelihood deviance; (b) plot of LASSO coefficient profiles. Each colorful curve represents the LASSO coefficient profile of a feature against the lnλ sequence. The values above the figure represent the numbers of variables included in the model, given the corresponding λ shown on the *x*-axis.

3.4. Association of IgA, CRP, CK, and age with improved outcomes

Serum levels of IgA, CRP, and CK were divided into quartiles as categorical variables. The median and proportion of improved patients in each quartile are presented in Table 4. Compared with the first quartile of IgA (reference), the probability of improvement decreased by the quartile of IgA level: The odds ratios (ORs) were 0.37 (95% CI, 0.07–1.54) for the second quartile, 0.25 (95% CI, 0.05–0.97) for the third quartile, and 0.20 (95% CI, 0.04–0.76) for the fourth quartile. Significant results from the trend test also confirmed the relationship between IgA levels and improved outcomes. Similar results were obtained from the performance of the same analyses on CRP and CK levels, as shown in Table 4. In addition, OR was 0.032 (95% CI, 0.001–0.564) for the $\geq$ 70 years of age group compared with the youngest group, suggesting that the elderly had greater difficulty recovering from the illness. The trend examination showed an association between increasing age and a reduction in the likelihood of prognosis improvement, although no significant effects on disease relief were observed in the second and third age groups compared with the first age group (Table 5).

3.5. Clinical utility of a nomogram

Based on the results of the multivariate logistics regression analyses, we further constructed a nomogram by combining

Table 3
Predictive variables for the probability of improved outcomes.

Variable	Prediction model		
	Regression coefficient	OR (95% CI)	*P* value
IgA	−0.011	0.989 (0.979–0.998)	0.040
CRP	−0.034	0.966 (0.942–0.986)	0.003
APACHE II	−1.716	0.180 (0.045–0.575)	0.006
CK	−2.807	0.006 (0.001–0.437)	0.008
APACHE II : CK	0.188	1.207 (1.021–1.439)	0.022

OR: odds ratio; CI: confidence interval; APACHE II : CK: the interaction between APACHE II and CK.

Table 4
Association of ascending quartiles of IgA, CRP, and CK levels with improved outcomes.

Variable	Quartile of serum levels				*P* for trend
	1	2	3	4	
IgA					
Median (mg·dL^{-1})	140	192	235	311	
Improved number	23 (88.5%)	20 (76.9%)	17 (65.4%)	15 (57.7%)	
OR (95% CI)	1.00 (reference)	0.37 (0.07–1.54)	0.25 (0.05–0.97)	0.20 (0.04–0.76)	0.038
CRP					
Median (mg·L^{-1})	3.4	15.4	34.6	87.4	
Improved number	24 (92.3%)	23 (88.5%)	18 (69.2%)	10 (38.5%)	
OR (95% CI)	1.00 (reference)	0.64 (0.08–4.19)	0.19 (0.03–0.86)	0.05 (0.01–0.23)	< 0.001
CK					
Median (U·L^{-1})	40	60	92	199	
Improved number	22 (84.6%)	20 (76.9%)	19 (73.1%)	14 (53.8%)	
OR (95% CI)	1.00 (reference)	0.42 (0.07–2.35)	0.26 (0.05–0.81)	0.19 (0.03–0.72)	0.036

Note: Test for trend based on variable containing median value for each quartile.

Table 5
Correlation between increasing age and improved outcomes.

Age (years)	Patients	Improved number	Unadjusted OR (95% CI)	*P* value	Adjusted OR (95% CI)[a]	*P* value
< 40	25	23 (92.0%)	1.000 (reference)		1.000 (reference)	
40–54	27	22 (81.5%)	0.383 (0.051–1.981)	0.279	0.301 (0.007–6.832)	0.471
55–69	35	27 (77.1%)	0.293 (0.041–1.315)	0.144	0.161 (0.005–2.875)	0.247
≥ 70	17	3 (17.6%)	0.019 (0.002–0.103)	< 0.001	0.032 (0.001–0.564)	0.030
P for trend			< 0.001		0.018	

[a] Adjusted for the variables included in the final model, as shown in Table 3.

prognostic factors including IgA, CRP, CK, APACHE II, and the interaction between CK and APACHE II. A quantitative method was made accessible for clinicians to predict the probability of improved prognosis in each COVID-19 patient (Fig. 2). Each patient is given a point for each prognostic parameter, and the distribution of the score is shown in a density plot. The higher the total number of points, the more likely the patient is to improve. Moreover, calibration curves demonstrated that the nomogram had a similar performance compared with the ideal model. The apparent curve confirmed the good prediction capability of our nomogram (Fig. 3). In addition, the decision curve showed that making use of this nomogram for predicting the probability of improved prognosis would gain more net benefits than an all-or-none patient intervention scheme if the threshold probability was less than 88%, which suggests a high potential for clinical application (Fig. 4). Stratification of the improvement probability for 1000 samples was predicted on the clinical impact curve (Fig. 5). The predictive improved number was close to the actual number of positive cases when the threshold probability was greater than 0.2. At this time, the cost-to-benefit ratio was 0.25.

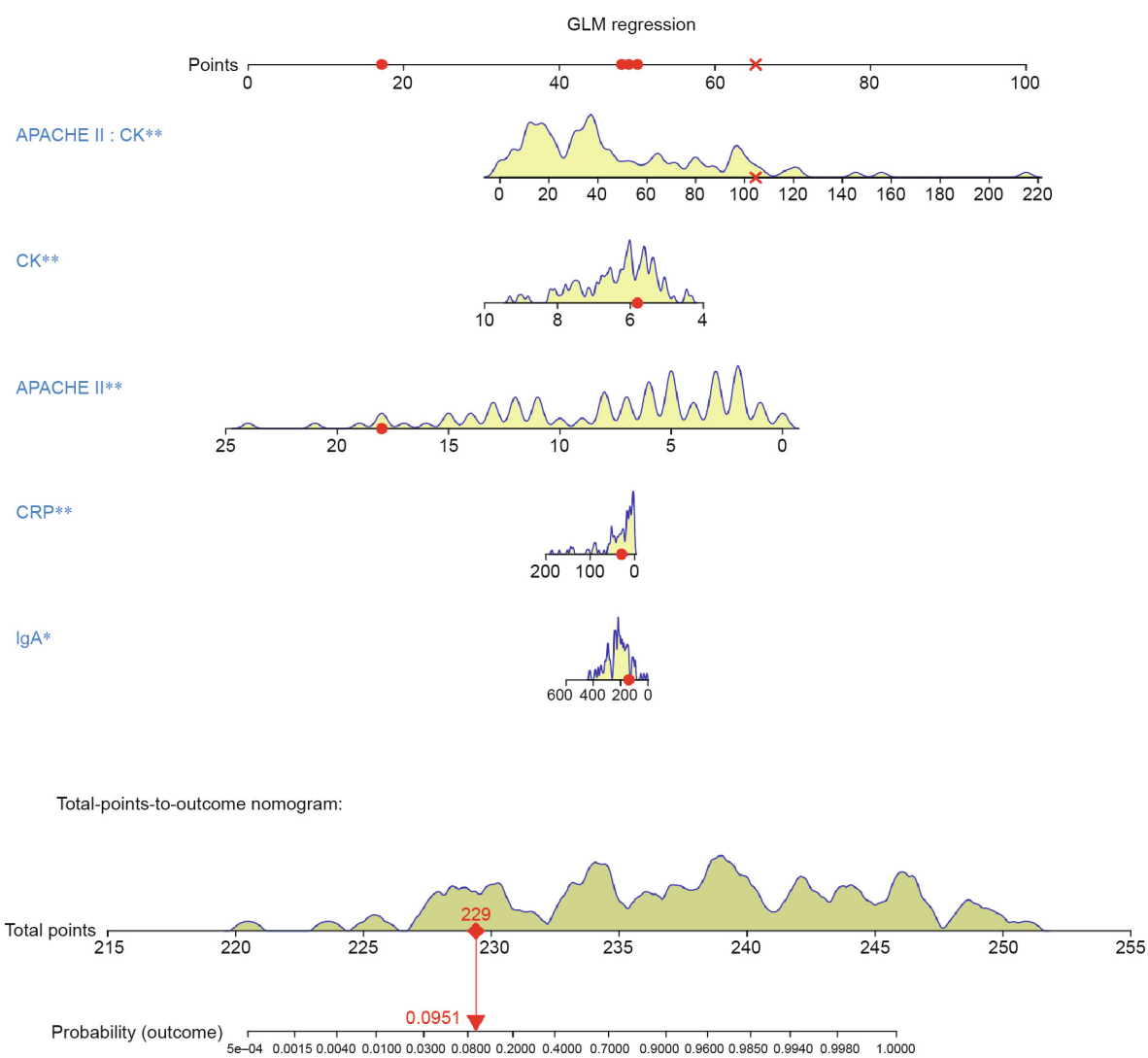

Fig. 2. Nomogram to predict probability of improved outcomes in COVID-19 patients. Yellow density plots describe the distribution of COVID-19 patients in prognostic parameters and total points. Yellow density plots describe the distribution of COVID-19 patients in prognostic parameters and total points while the red dots and cross represent one patient's points as an example. GLM: general multivariate regression. *: $P \leq 0.05$, **: $P \leq 0.01$.

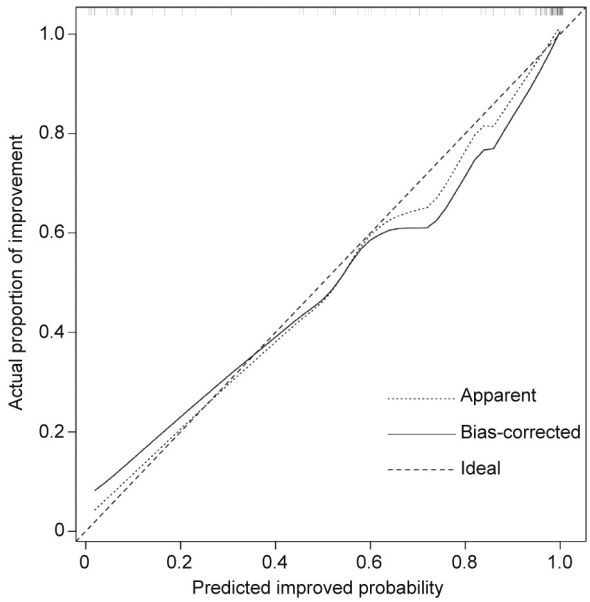

Fig. 3. Calibration curve for the nomogram to predict probability of improved outcomes. The *x*-axis represents the predicted improved probability and the *y*-axis denotes the actual proportion of improvement. The diagonal dotted line indicates the best prediction by an ideal model. The apparent line represents the uncorrected performance of the nomogram while the solid line shows the bais-corrected performance. 1000 bootstrap repetitions; mean absolute error = 0.029; n = 104.

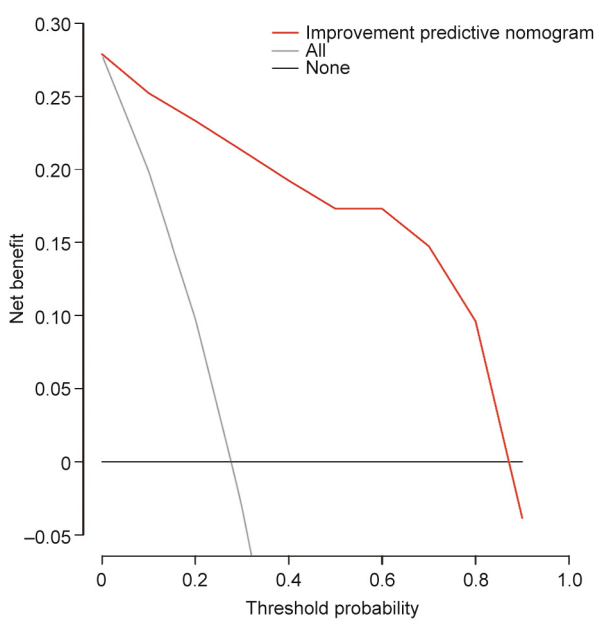

Fig. 4. Decision curve for the improvement predictive nomogram. The net benefits were measured at different threshold probabilities. The red line represents the improvement predictive nomogram. The gray line represents the assumption that all patients have improved outcomes. The black line represents the assumption that no patients have improved outcomes.

4. Discussion

Despite worldwide efforts to contain the new coronavirus, hotspots continue to emerge, and the number of cases is on the rise. As of 2 March 2020, the SARS-CoV-2 has infected over 90 900 people and killed 3118 [10]. Although recently published articles have reported the clinical, virological, and epidemiological characteristics of patients with COVID-19 [5,11], few studies have focused on prognostic indicators or risk factors. Thus, we constructed this

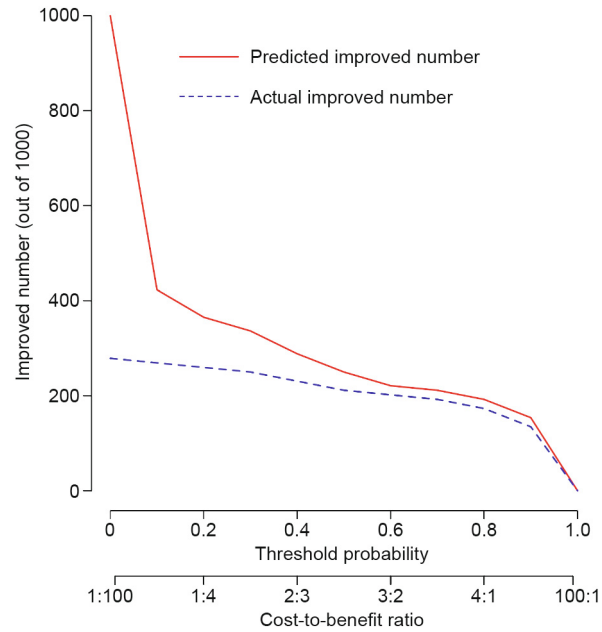

Fig. 5. Clinical impact curve to predict the improved number for a population size of 1000. The red curve shows the predicted improved number at different threshold probabilities and the blue curve represents actual improved patients.

predictive nomogram using individual factors to make accurate prognostic assessments in order to quantitatively predict clinical outcomes in a personalized way. This is an urgent, user-friendly, and easy-to-use method.

We reported on 104 patients, of which 75 had improved outcomes and 29 did not. The nomogram established in this study suggested five prognostic factors for predicting the outcome: APACHE II, CK, CRP, IgA, and the interaction between CK and APACHE II. Similar to previous findings for 51 patients with MERS-CoV infection [16], we found that the widely used disease classification system APACHE II in the ICU [14] was associated with the prognosis, with higher APACHE II scores leading to worse outcome. The APACHE II score was calculated using the acute physiology score (APS), chronic physiology score (CPS), and age. Several factors included in the APS showed significant differences between improved patients and those without improvement, according to our results. Vital signs and laboratory parameters have also shown significant differences between ICU and non-ICU patients with COVID-19 [11]. Moreover, researchers have reported comorbidities and age as risk factors of severity and mortality in patients with SARS-CoV [17,23] and MERS-CoV infection [18–22].

In addition, CK was at a higher level for patients admitted to the ICU [11], which was consistent with our model's prediction. This finding may be attributed to muscle damage caused by COVID-19, similar to changes in SARS [23]. While muscle weakness and elevated levels of serum CK occurred in more than 30% of the SARS-infected patients, focal myofiber necrosis was observed in a series of postmortem cases [23]. As for COVID-19, the first autopsy revealed a gray-red fish-shaped myocardial section. However, it remains uncertain whether this myocardial damage was due to an original heart disease or a viral infection, and further research is needed. We speculate that myopathy is also likely to play an important role in COVID-19. In patients with MERS-CoV, CRP is a common predictor of the development of pneumonia and respiratory failure associated with thrombocytopenia and lymphocytopenia [24]. Similarly, CRP may be related to enhanced inflammation and cytokine storms caused by COVID-19 invasion.

Interestingly, although IgA has been acknowledged as the first barrier against the virus in the respiratory tract due to the mucosal

immune system, a higher IgA level led to worse outcomes based on our findings [25]. This might result from the fact that the IgA we measured was from blood, rather than secretory IgA (sIgA) from the mucus. Unlike sIgA, serum IgA can cause antibody-dependent cell-mediated cytotoxicity (ADCC), lead to degranulation of eosinophils and basophils, result in phagocytosis by monocytes, macrophages, and neutrophils, and trigger respiratory burst activity by polymorphonuclear leukocytes [26], which may be related to sustained inflammatory response and cytokine storm. The three pathological mechanisms associated with the prognostic factors selected in our study—namely, sustained inflammatory response, cytokine storm, and direct effects of the virus—are likely to have negative effects on outcome improvement.

Thus far, a specific treatment method for coronavirus infection has not been found. It is important to identify risk factors that can predict and improve patient prognosis through personalized treatment methods. Therefore, we constructed this nomogram to quantitatively measure the severity of infected patients and predict the subsequent outcomes of infected patients. For high-risk patients, early use of high-flow oxygen therapy, non-invasive ventilation, or even invasive ventilation is recommended.

5. Limitations of this study

First, the number of patients in this study limits further enhancement of the predictive power of our nomogram. Second, we could not determine the final outcome of some patients because their condition was still changing as of the study submission. Third, all patients were admitted to our hospital in Zhejiang Province, which likely resulted in regional limitations. This predictive nomogram requires further validation at different centers in the future.

Acknowledgments

This work was supported by the research on the prevention and clinical treatment in patients with COVID-19 (2020C03123), a funding of the Zhejiang Provincial Department of Science and Technology; the National Natural Science Foundation of China (81790631); and the National Key Research and Development Program of China (2018YFC2000500).

Ethical approval

This study was approved by the Ethics Committee of the First Affiliated Hospital, Collegue of Medicine, Zhejiang University (2020IIT A0040).

Compliance with ethics guidelines

Jiaojiao Xie, Ding Shi, Mingyang Bao, Xiaoyi Hu, Wenrui Wu, Jifang Sheng, Kaijin Xu, Qing Wang, Jingjing Wu, Kaicen Wang, Daiqiong Fang, Yating Li, and Lanjuan Li declare that they have no conflict of interest or financial conflicts to disclose.

References

[1] Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet 2020;395(10223):497–506.
[2] Chen N, Zhou M, Dong X, Qu J, Gong F, Han Y, et al. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. Lancet 2020;395(10223):507–13.
[3] World Health Organization. Naming the coronavirus disease (COVID-19) and the virus that causes it [Internet]. Geneva: Word Health Organization; c2020 [cited 2020 Feb 11]. Available from: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/naming-the-coronavirus-disease-(covid-2019)-and-the-virus-that-causes-it.
[4] Holshue ML, DeBolt C, Lindquist S, Lofy KH, Wiesman J, Bruce H, et al. First case of 2019 novel coronavirus in the United States. N Engl J Med 2020;382(10):929–36.
[5] Xu X, Wu X, Jiang X, Xu K, Ying L, Ma C, et al. Clinical findings in a group of patients infected with the 2019 novel coronavirus (SARS-Cov-2) outside of Wuhan, China: retrospective case series. BMJ 2020;368:m606.
[6] Zaki AM, van Boheemen S, Bestebroer TM, Osterhaus AD, Fouchier RA. Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. N Engl J Med 2012;367:1814–20.
[7] Zhong N, Zheng B, Li Y, Poon, Xie Z, Chan K, et al. Epidemiology and cause of severe acute respiratory syndrome (SARS) in Guangdong, People's Republic of China, in February, 2003. Lancet 2003;362(9393):1353–8.
[8] Guan W, Ni Z, Hu Y, Liang W, Qu C, He J, et al. Clinical characteristics of coronavirus disease 2019 in China. N Engl J Med 2019;2020(382):1708–20.
[9] nhc.gov.cn/[Internet]. Beijing: National Health Commission of the People's Republic of China; [cited 2020 Mar 3]. Available from: http://www.nhc.gov.cn. Chinese.
[10] Coronavirus Disease (COVID-19) Dashboard [Internet]. Geneva: Word Health Orgainization; c2020 [cited 2020 Mar 2]. Available from: https://covid19.who.int.
[11] Wang D, Hu B, Hu C, Zhu F, Liu X, Zhang J, et al. Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus-infected pneumonia in Wuhan, China. JAMA 2020;323(11):1061–9.
[12] Peng J, Wang M, Ang IYH, Tan SHX, Lewis RF, Chen J, et al. Potential rapid diagnostics, vaccine and therapeutics for 2019 novel coronavirus (2019-nCoV): a systematic review. J Clin Med 2020;9(3):623.
[13] Clinical management of severe acute respiratory infection when COVID-19 is suspected: interim guidance [Internet]. Geneva: World Health Organization; c2020 [cited 2020 Mar 13]. Available from: https://www.who.int/publications-detail/clinical-management-of-severe-acute-respiratory-infection-when-novel-coronavirus-(ncov)-infection-is-suspected.
[14] Knaus WA, Draper EA, Wagner DP, Zimmerman JE. APACHE II: a severity of disease classification system. Crit Care Med 1985;13(10):818–29.
[15] National Health Commission of the People's Republic of China; National Administration of Traditional Chinese Medicine. Guidelines for diagnosis and treatment of novel coronavirus pneumonia (trial version 7) [Internet]. Beijing: The State Council of the People's Republic of China; 2020 Mar 3 [cited 2020 Mar 4]. Available from: http://www.gov.cn/zhengce/zhengceku/2020-03/04/5486705/files/ae61004f930d47598711a0d4cbf874a9.pdf. Chinese.
[16] Al Ghamdi M, Alghamdi KM, Ghandoora Y, Alzahrani A, Salah F, Alsulami A, et al. Treatment outcomes for patients with Middle Eastern respiratory syndrome coronavirus (MERS CoV) infection at a coronavirus referral center in the Kingdom of Saudi Arabia. BMC Infect Dis 2016;16:174.
[17] Wu J, Xu F, Zhou W, Feikin DR, Lin C, He X, et al. Risk factors for SARS among persons without known contact with SARS patients, Beijing, China. Emerg Infect Dis 2004;10(2):210–6.
[18] Saad M, Omrani AS, Baig K, Bahloul A, Elzein F, Matin MA, et al. Clinical aspects and outcomes of 70 patients with Middle East respiratory syndrome coronavirus infection: a single-center experience in Saudi Arabia. Int J Infect Dis 2014;29:301–6.
[19] Feikin DR, Alraddadi B, Qutub M, Shabouni O, Curns A, Oboho IK, et al. Association of higher MERS-CoV virus load with severe disease and death, Saudi Arabia, 2014. Emerg Infect Dis 2015;21:2029–35.
[20] Majumder MS, Kluberg SA, Mekaru SR, Brownstein JS. Mortality risk factors for Middle East respiratory syndrome outbreak, South Korea, 2015. Emerg Infect Dis 2015;21(11):2088–90.
[21] Ahmed AE. The predictors of 3- and 30-day mortality in 660 MERS-CoV patients. BMC Infect Dis 2017;17:615.
[22] Nam HS, Park JW, Ki M, Yeon MY, Kim J, Kim SW. High fatality rates and associated factors in two hospital outbreaks of MERS in Daejeon, the Republic of Korea. Int J Infect Dis 2017;58:37–42.
[23] Leung TW, Wong KS, Hui AC, To KF, Lai ST, Ng WF, et al. Myopathic changes associated with severe acute respiratory syndrome: a postmortem case series. Arch Neurol 2005;62(7):1113–7.
[24] Ko JH, Park GE, Lee JY, Lee JY, Cho SY, Ha YE, et al. Predictive factors for pneumonia development and progression to respiratory failure in MERS-CoV infected patients. J Infect 2016;73(5):468–75.
[25] Liew FY, Russell SM, Appleyard G, Brand CM, Beale J. Cross-protection in mice infected with influenza A virus by the respiratory route is correlated with local IgA antibody rather than serum antibody or cytotoxic T cell reactivity. Eur J Immunol 1984;14(4):350–6.
[26] Otten MA, van Egmond M. The Fc receptor for IgA (FcalphaRI, CD89). Immunol Lett 2004;92(1–2):23–31.