



OPEN

## A risk scoring system to predict progression to severe pneumonia in patients with Covid-19

Ji Yeon Lee<sup>1,2</sup>, Byung-Ho Nam<sup>3</sup>, Mhinjine Kim<sup>2,4</sup>, Jongmin Hwang<sup>5</sup>, Jin Young Kim<sup>6</sup>, Miri Hyun<sup>1,2</sup>, Hyun Ah Kim<sup>1,2</sup> & Chi-Heum Cho<sup>1b,2,7✉</sup>

Rapid outbreak of coronavirus disease 2019 (Covid-19) raised major concern regarding medical resource constraints. We constructed and validated a scoring system for early prediction of progression to severe pneumonia in patients with Covid-19. A total of 561 patients from a Covid-19 designated hospital in Daegu, South Korea were randomly divided into two cohorts: development cohort (N = 421) and validation cohort (N = 140). We used multivariate logistic regression to identify four independent risk predictors for progression to severe pneumonia and constructed a risk scoring system by giving each factor a number of scores corresponding to its regression coefficient. We calculated risk scores for each patient and defined two groups: low risk (0 to 8 points) and high risk (9 to 20 points). In the development cohort, the sensitivity and specificity were 83.8% and 78.9%. In the validation cohort, the sensitivity and specificity were 70.8% and 79.3%, respectively. The C-statistics was 0.884 (95% CI 0.833–0.934) in the development cohort and 0.828 (95% CI 0.733–0.923) in the validation cohort. This risk scoring system is useful to identify high-risk group for progression to severe pneumonia in Covid-19 patients and can prevent unnecessary overuse of medical care in limited-resource settings.

After the Coronavirus disease 2019 (Covid-19) first occurred in China on December 2019, it has become pandemic in March<sup>1</sup>. Due to the surge in Covid-19 patients, lack of medical resources (e.g., negative pressure rooms, personal protective equipment and medical personnel) has become a major problem. During the height of the outbreak in early March, the Ministry of Health and Korea Centers for Disease Control and Prevention established a policy guideline of assigning patients with Covid-19 to different treatment centers based on the level of severity and risk in order to maximize the usage of medical resources. Patients with mild symptoms were assigned to community treatment centers to recover whereas high-risk patients with old age or moderate to severe cases were hospitalized. However, there has been an increasing number of patients classified as mild cases at admission who have developed severe pneumonia during hospitalization.

It is important to provide timely critical care to Covid-19 patients with serious symptoms to reduce number of deaths and burden on overall health system<sup>2</sup>. Therefore, it is key to identify patients who are at higher risk of developing severe pneumonia in early stages of disease. Previous studies show that established and potential risk factors associated with Covid-19 complications are older age (e.g., > 65 years), cardiovascular disease, chronic lung disease, diabetes mellitus, obesity, immunocompromise, end-stage renal disease, and liver disease<sup>3–7</sup>. The purpose of this study is to present a novel scoring system that could be used by clinicians to predict progression to severe pneumonia in Covid-19 patients in earlier stages.

<sup>1</sup>Department of Infectious Diseases, Keimyung University Dongsan Hospital, Keimyung University School of Medicine, Daegu, Republic of Korea. <sup>2</sup>Covid-19 Task Force Team of Keimyung University Daegu Dongsan Hospital, Daegu, Republic of Korea. <sup>3</sup>HERINGS, Institute of Advanced Clinical and Biomedical Research, Seoul, Republic of Korea. <sup>4</sup>Division of Health Policy and Administration, School of Public Health, University of Illinois at Chicago, Chicago, USA. <sup>5</sup>Department of Cardiology, Keimyung University Dongsan Hospital, Keimyung University School of Medicine, Daegu, Republic of Korea. <sup>6</sup>Department of Radiology, Keimyung University Dongsan Hospital, Keimyung University School of Medicine, Daegu, Republic of Korea. <sup>7</sup>Department of Obstetrics and Gynecology, Keimyung University Dongsan Hospital, Keimyung University School of Medicine, 1035, Dalgubeol-daero, Dalseo-gu, Daegu 42601, Republic of Korea. ✉email: chcho@kmu.ac.kr

## Methods

**Study population.** Keimyung University Daegu Dongsan Hospital (KDDH) were designated as Covid-19 regional hospital on February 21, 2020 during a large outbreak of Covid-19 in Daegu, South Korea. We enrolled all patients ( $\geq 18$  years old) with confirmed Covid-19 admitted to this hospital from February 21, 2020 to March 31, 2020. Patients presenting with severe Covid-19 pneumonia at admission were excluded. A Covid-19 patient was defined as an individual with laboratory confirmation of Covid-19 which required positive results of SARS-CoV-2 RNA, irrespective of clinical signs and symptoms. Severe Covid-19 pneumonia was defined as at least one of the followings: (1) resting oxygen saturation  $\leq 93\%$  in room air, or (2)  $\text{PaO}_2/\text{FiO}_2 \leq 300$  mmHg or requirement of mechanical ventilation<sup>8</sup>. Patients were classified into two groups: Progression (progression to severe pneumonia during hospitalization) and Stable (not requiring oxygen treatment during hospitalization). All patients with stable Covid-19 infection during hospitalization were followed for more than 4 weeks after admission. The study was conducted in accordance with the Declaration of Helsinki. Informed written consent was obtained from all participants. This study was approved by the Institutional Review Board at the Keimyung University Dongsan Hospital (IRB No. 2020-03-027).

**Data collection.** The presenting history, comorbidity status, epidemiologic history, and vital signs were collected. Comorbidities were defined as having at least one of the followings: hypertension, diabetes mellitus, hyperlipidemia, cardiovascular disease, chronic pulmonary disease, chronic liver disease, chronic kidney disease, neurologic disease, autoimmune disease, malignancy and HIV infections. The laboratory parameters, including complete blood count, liver and renal function, C-reactive protein (CRP), lactate dehydrogenase (LDH), creatinine phosphokinase (CPK) were examined at admission. The oxygen saturation was measured by pulse oxygen saturation on room air at rest state. All patients were examined by chest X-ray. Clinical outcome (progression of illness) were monitored up to April 26, 2020. Data were collected retrospectively by electronic medical chart review.

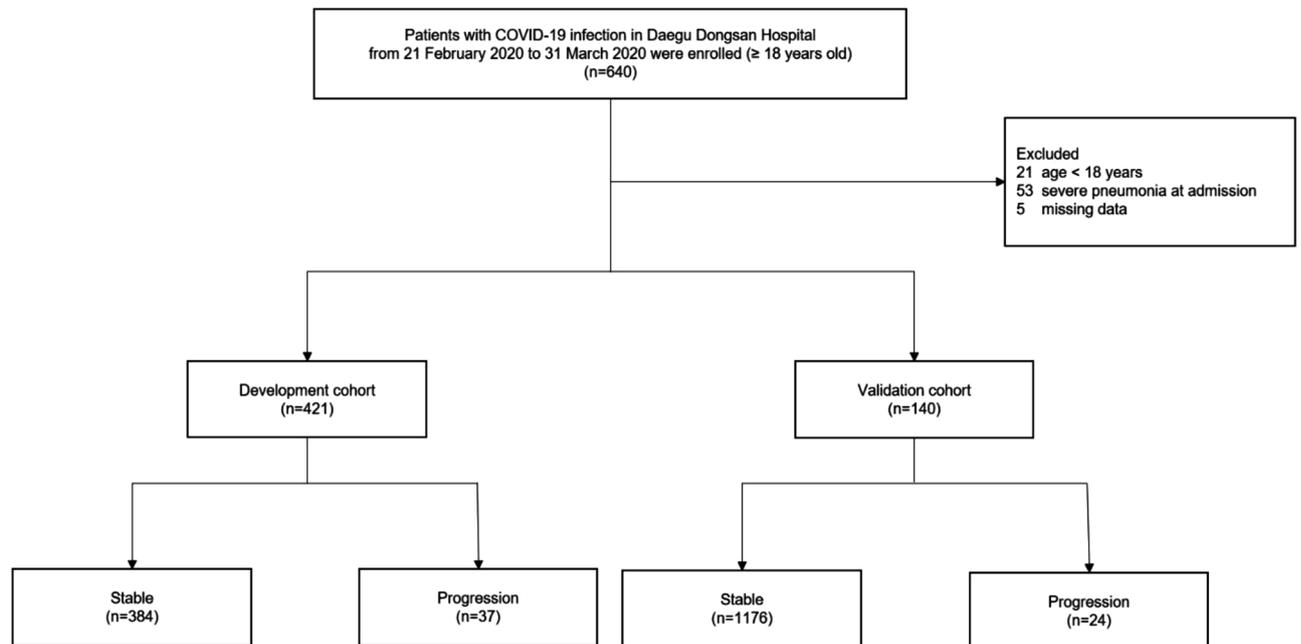
**Statistical analysis.** Potential risk factors for progression to severe pneumonia were evaluated. The Wilcoxon rank sum test was used for continuous variables and the Chi-square test was used for categorical variables to examine the baseline difference between stable and progressed patient group. Most continuous variables that showed significant difference at the baseline were dichotomized. The cut-off values for each continuous variable were determined such that it showed the best discriminatory ability based on the Youden index (sensitivity + specificity - 1). Age group and LDH group were categorized based on the risk probability trend of specific sections within age (by 5 years) and LDH (50 U/L). Age-adjusted logistic regression model was used to identify risk factors. Multivariate logistic regression model with stepwise selection process was used to develop a risk prediction model. *P* value of 0.05 was used for variable selection process. The model performance was evaluated with respect to its discrimination and calibration ability. Discrimination was quantified using the C-statistics and Hosmer–Lemeshow (H–L)  $\chi^2$  statistic was calculated for calibration. The Receiver Operating Characteristics (ROC) Curve for the C-statistics was generated and the square distance between observed prevalence and mean predicted probabilities for each quintile of predicted risk was assessed for calibration. The model performance was also evaluated on the separate validation cohort. Statistical analyses were performed using SAS version 9.4 (SAS institute, Cary, NC), and R package version 3.6.2.

## Results

**Clinical characteristics of patients.** The selection of the study population is illustrated in Fig. 1. A total 640 patients were hospitalized from February 21 through March 31, 2020, the follow-up period ended in April 26, 2020. 21 cases were excluded for the age under 18 years old, 53 patients were excluded for severe pneumonia at admission, and 5 patients were excluded for lack of laboratory examination. Out of 561 patients, 421 were randomly placed in the development cohort and 140 were assigned to the validation cohort, respectively. All patients with stable Covid-19 during hospitalization were followed for more than 4 weeks after admission. Demographics and clinical characteristics of patients who remained stable and who had progressed to severe pneumonia in Covid-19 infection are summarized in Supplementary Table 1 for the development cohort and Supplementary Table 2 for the validation cohort. In the development cohort, 37 (8.8%) of 421 patients had progressed to severe pneumonia. In the validation cohort, 24 patients (17.1%) developed severe Covid-19 pneumonia.

**Independent risk factors associated with progression to severe pneumonia.** A risk prediction model was developed in the development cohort. In the univariate analysis, age, male gender, comorbidities, initial chest X-ray abnormality, absolute neutrophil count (ANC), absolute lymphocyte count (ALC), platelet count, blood urea nitrogen (BUN), estimated glomerular filtration rate (eGFR), aspartate aminotransferase (AST), albumin, CRP, CPK and LDH were significantly associated with progression to severe pneumonia.

The age-adjusted univariate analysis revealed male gender, comorbidities, initial chest X-ray abnormality, ALC, platelet, BUN, AST, albumin, CRP, CPK and LDH as significantly associated factors. Age, hemoglobin, CRP and LDH were included in the final model by stepwise selection process in the multivariate logistic regression model (Table 1). In the development cohort, the proportions of progression to severe pneumonia in each category were seen as follows: age [ $< 50$  years (0.7%), 50–59 years (6.0%), 60–69 years (10.3%), 70–79 years (17.6%),  $\geq 80$  years (37.5%)], hemoglobin [ $< 13$  g/L (7.5%),  $\geq 13$  g/L (11.3%)], CRP [ $< 1.4$  mg/dL (4.0%),  $\geq 1.4$  mg/dL (26.1%)], LDH [ $< 500$  U/L (3.9%), 500–700 U/L (15.1%),  $\geq 700$  U/L (45.8%)]. In the validation cohort, the respective proportions were: age [ $< 50$  years (6.5%), 50–59 years (16.7%), 60–69 years (26.7%), 70–79 years (13.6%),  $\geq 80$  years (66.7%)], hemoglobin [ $< 13$  g/L (13.7%),  $\geq 13$  g/L (24.4%)], CRP [ $< 1.4$  mg/dL (8.3%),  $\geq 1.4$  mg/dL (46.9%)], LDH [ $< 500$  U/L (10.6%), 500–700 U/L (18.2%),  $\geq 700$  U/L (61.5%)] (Supplementary Fig. 1).



**Figure 1.** Selection of study patients.

**Construction and performance of KDDH scoring system.** A scoring system was developed such that the points given to each variable correspond to its regression coefficients in the model (Table 2). The maximum points that a patient can get is 20. This scoring system was constructed based on the four factors, and two patient groups were constructed based on risk scores: low risk (0 to 8 points) and high risk (9 to 20 points). We identified patient group with over the score of 9 to have higher possibility of developing severe pneumonia. Sensitivity and specificity in the development cohort were 83.8% and 78.9%, respectively. In the development cohort, 31 (27.7%) of high-risk group and 6 (1.9%) of low-risk group progressed to severe pneumonia.

The discrimination ability was evaluated for both prediction model and score system in terms of C-statistics. The C-statistics for the model was 0.886 (95% CI 0.836–0.937) and 0.884 (95% CI 0.833–0.934) for the score system (Fig. 2).

**Validation of the risk prediction models.** In the validation cohort, 17 (41.5%) of high-risk group and 7 (7.1%) of low-risk group developed severe pneumonia. Baseline characteristics for the validation cohort were similar to those in the development cohort (Supplementary Table 2). At the cut-off value of 9 points for high-risk group, sensitivity and specificity of score system in the development cohort were 70.8% and 79.3%, respectively (Table 3). The C-statistics were 0.826 (95% CI 0.730–0.922) for the fitted value model and 0.828 (95% CI 0.733–0.923) for the scoring system in the validation cohort (Fig. 2). The model calibration on the validation cohort are shown in the Fig. 3. There was no significant difference between predicted progression to severe pneumonia with KDDH score and actual progression ( $\chi^2 = 2.448$ ,  $p$  value = 0.485).

## Discussion

This study presented a novel scoring system which can be used to predict progression of Covid-19. Identifying Covid-19 patients who are at a higher risk of developing severe symptoms in early stages can inform better medical resource allocation and patient care during massive outbreaks. We found age, CRP, LDH and hemoglobin as independent high-risk factors associated with progression of Covid-19 infection. Some of our findings are consistent with previous studies that identified different risk factors to be associated with poor clinical outcomes in patients with Covid-19<sup>6,7,9,10</sup>. Older age, high Sequential Organ Failure Assessment (SOFA) score, and d-dimer greater than 1  $\mu\text{g}/\text{mL}$  at admission were described as potential risk factors for mortality of in-hospital patients with Covid-19<sup>6</sup>. Wang et al.<sup>9</sup> reported older age, dyspnea, lymphopenia, comorbidities (e.g., cardiovascular disease), and acute respiratory distress syndrome (ARDS) as predictors of fatal outcomes in elderly Covid-19 patients. Age, sex, CRP, LDH, lymphocyte count, and features derived from CT images were most reported predictors of severe Covid-19 progression<sup>10</sup>. Jiang X et al. presented research on progression to ARDS through AI framework, and it mentioned higher hemoglobin as one of the risk factors of later development of ARDS. Moreover, higher hemoglobin levels were associated male gender or even unreported tobacco use<sup>11</sup>.

To our knowledge, KDDH scoring system is one of few scoring systems to predict Covid-19 progression at early stage. Ji et al.<sup>12</sup> created a CALL score model with age, comorbidities, lymphopenia and LDH to predict Covid-19 progression at early stage. This is a relatively simple scoring system that only require basic tests for laboratory parameters. However, it has limitations of smaller sample size and no validation. There exists difference in definition of progression to severe Covid-19. CALL model defines deteriorated chest radiologic findings

| Risk factor                      | Univariate analysis     |         | Age-adjusted univariable analysis |         | Multiple analysis   |         |
|----------------------------------|-------------------------|---------|-----------------------------------|---------|---------------------|---------|
|                                  | OR (95% CI)             | p value | OR (95% CI)                       | p value | OR (95% CI)         | p value |
| Age, years                       |                         |         |                                   |         |                     |         |
| < 50                             | 1.00 (Ref.)             |         |                                   |         |                     |         |
| 50–59                            | 9.00 (1.51–171.42)      | 0.043   |                                   |         | 6.63 (0.75–58.96)   | 0.090   |
| 60–69                            | 16.27 (2.98–302.81)     | 0.009   |                                   |         | 7.93 (0.91–69.41)   | 0.061   |
| 70–79                            | 30.21 (5.75–556.93)     | 0.001   |                                   |         | 24.99 (2.91–214.99) | 0.003   |
| ≥ 80                             | 84.60 (14.50 – 1616.18) | < 0.001 |                                   |         | 69.28 (7.45–644.60) | < 0.001 |
| Gender, male                     |                         |         |                                   |         |                     |         |
| Female                           | 1.00 (Ref.)             |         | 1.00 (Ref.)                       |         |                     |         |
| Male                             | 3.08 (1.55–6.15)        | 0.001   | 3.20 (1.53–6.77)                  | 0.002   |                     |         |
| Comorbidities                    |                         |         |                                   |         |                     |         |
| Without                          | 1.00 (Ref.)             |         | 1.00 (Ref.)                       |         |                     |         |
| With                             | 3.70 (1.80–8.23)        | 0.001   | 1.33 (0.58–3.19)                  | 0.002   |                     |         |
| Initial CXR                      |                         |         |                                   |         |                     |         |
| Normal                           | 1.00 (Ref.)             |         | 1.00 (Ref.)                       |         |                     |         |
| Lung infiltration                | 6.99 (3.16–17.68)       | < 0.001 | 4.30 (1.87–11.26)                 | 0.001   |                     |         |
| ANC, / $\mu$ L                   |                         |         |                                   |         |                     |         |
| < 3010                           | 1.00 (Ref.)             |         | 1.00 (Ref.)                       |         |                     |         |
| ≥ 3010                           | 2.91 (1.46–6.06)        | 0.003   | 1.62 (0.76–3.56)                  | 0.218   |                     |         |
| ALC, / $\mu$ L                   |                         |         |                                   |         |                     |         |
| ≥ 1000                           | 1.00 (Ref.)             |         | 1.00 (Ref.)                       |         |                     |         |
| < 1000                           | 4.64 (2.10–9.88)        | < 0.001 | 3.14 (1.34–7.11)                  | 0.007   |                     |         |
| Hemoglobin, g/dL                 |                         |         |                                   |         |                     |         |
| < 13.3                           | 1.00 (Ref.)             |         | 1.00 (Ref.)                       |         |                     |         |
| ≥ 13.3                           | 1.58 (0.79–3.12)        | 0.191   | 2.95 (1.36–6.46)                  | 0.006   | 3.79 (1.30–7.31)    | 0.011   |
| Platelet, $10^3$ /uL             |                         |         |                                   |         |                     |         |
| ≥ 150                            | 1.00 (Ref.)             |         | 1.00 (Ref.)                       |         |                     |         |
| < 150                            | 2.84 (1.19–6.27)        | 0.013   | 3.08 (1.21–7.45)                  | 0.014   |                     |         |
| BUN, mg/dL                       |                         |         |                                   |         |                     |         |
| < 19                             | 1.00 (Ref.)             |         | 1.00 (Ref.)                       |         |                     |         |
| ≥ 19                             | 4.59 (2.16–9.48)        | < 0.001 | 2.35 (1.04–5.11)                  | 0.035   |                     |         |
| eGFR, mL/min/1.73 m <sup>2</sup> |                         |         |                                   |         |                     |         |
| ≥ 60                             | 1.00 (Ref.)             |         | 1.00 (Ref.)                       |         |                     |         |
| < 60                             | 5.74 (2.06–14.77)       | < 0.001 | 2.30 (0.76–6.43)                  | 0.121   |                     |         |
| AST, U/L                         |                         |         |                                   |         |                     |         |
| < 30                             | 1.00 (Ref.)             |         | 1.00 (Ref.)                       |         |                     |         |
| ≥ 30                             | 3.88 (1.92–7.80)        | < 0.001 | 4.20 (1.93–9.29)                  | < 0.001 |                     |         |
| ALT, U/L                         |                         |         |                                   |         |                     |         |
| < 19                             | 1.00 (Ref.)             |         | 1.00 (Ref.)                       |         |                     |         |
| ≥ 19                             | 1.48 (0.75–3.03)        | 0.268   | 1.93 (0.90–4.40)                  | 0.103   |                     |         |
| Albumin, g/dL                    |                         |         |                                   |         |                     |         |
| ≥ 4.2                            | 1.00 (Ref.)             |         | 1.00 (Ref.)                       |         |                     |         |
| < 4.2                            | 6.14 (2.55–18.25)       | < 0.001 | 3.21 (1.27–9.87)                  | 0.023   |                     |         |
| CRP, mg/dL                       |                         |         |                                   |         |                     |         |
| < 1.4                            | 1.00 (Ref.)             |         | 1.00 (Ref.)                       |         |                     |         |
| ≥ 1.4                            | 8.58 (4.22–18.17)       | < 0.001 | 6.17 (2.88–13.78)                 | < 0.001 | 3.79 (1.52–9.44)    | 0.004   |
| CPK, U/L                         |                         |         |                                   |         |                     |         |
| < 69                             | 1.00 (Ref.)             |         | 1.00 (Ref.)                       |         |                     |         |
| ≥ 69                             | 2.31 (1.17–18.17)       | < 0.001 | 2.15 (1.04–4.50)                  | 0.039   |                     |         |
| LDH, U/L                         |                         |         |                                   |         |                     |         |
| < 500                            | 1.00 (Ref.)             |         | 1.00 (Ref.)                       |         |                     |         |
| 500–700                          | 4.31 (1.92–9.85)        | < 0.001 | 3.34 (1.41–8.07)                  | 0.006   | 2.12 (0.82–5.48)    | 0.119   |
| ≥ 700                            | 20.59 (7.68–56.45)      | < 0.001 | 17.33 (5.89–52.93)                | < 0.001 | 7.21 (2.12–24.56)   | 0.002   |

**Table 1.** Univariate and multiple logistic regression analysis in the development cohort. ALC, absolute lymphocyte count; ALT, Alanine aminotransferase; ANC, absolute neutrophil count; CI, confidence interval; AST, Aspartate aminotransferase; BUN, blood urea nitrogen; CPK, Creatinine phosphokinase; CRP, C-reactive protein; CXR, chest X-ray; eGFR, estimated glomerular filtration rate; LDH, lactate dehydrogenase; OR, odds ratio.

| Parameter                  | Points |
|----------------------------|--------|
| Age, years                 |        |
| < 50                       | 0      |
| 50–59                      | 4      |
| 60–69                      | 5      |
| 70–79                      | 7      |
| ≥ 80                       | 10     |
| C-reactive protein, mg/dL  |        |
| < 1.4                      | 0      |
| ≥ 1.4                      | 3      |
| Lactate dehydrogenase, U/L |        |
| < 500                      | 0      |
| 500–700                    | 2      |
| ≥ 700                      | 4      |
| Hemoglobin, g/dL           |        |
| < 13.3                     | 0      |
| ≥ 13.3                     | 3      |

**Table 2.** The calculator of KDDH score.

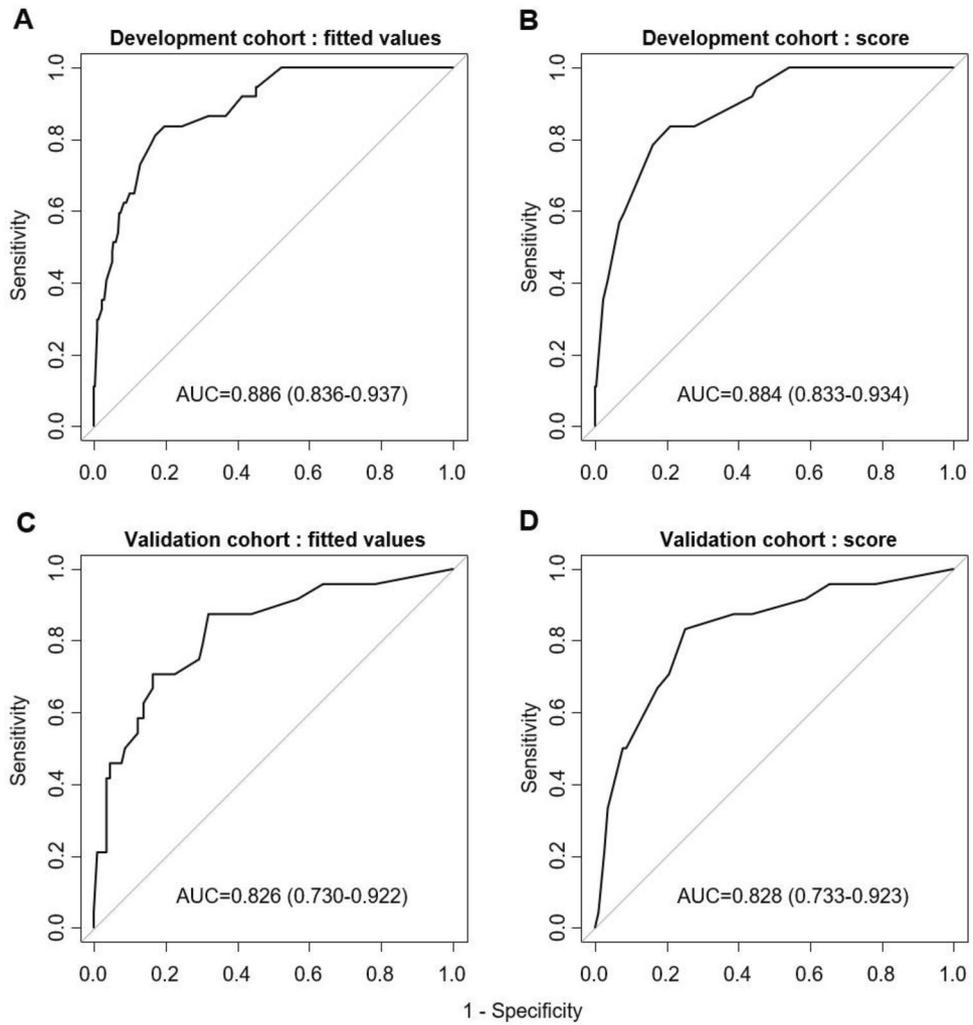
in progression, but the present study only defines patients requiring oxygen therapy as progression<sup>12</sup>. We did not include progression of radiologic findings without hypoxemia in the progression group, as deterioration of radiologic finding tends to be reflected later than clinical course. Gong et al.<sup>13</sup> presented a prognostic nomogram based on seven factors (older age, higher LDH and CRP, direct bilirubin, red blood cell distribution width (RDW), BUN, and lower albumin) to identify patients who are likely to develop severe Covid-19 infection. The nomogram was conducted with multicenter patients and was validated. However, its scoring system shows limited applicability due to score model complexity and greater number of clinical parameters. Liang et al.<sup>14</sup> developed a clinical risk score to predict risk of developing critical illness in Covid-19 patients, which was not associated with progression to severe pneumonia.

The KDDH scoring system has advantages of high sensitivity and specificity together with strong calibration that ensures no statistical difference between quintile of predicted risk. Moreover, this is uncomplicated and can predict progression to severe pneumonia by using a simple blood test that can be conducted in outpatient clinics.

High negative predictive value (98.1% in development cohort & 92.9% in validation cohort) of the presented scoring system demonstrates efficacy in identifying low risk patients, who can be managed at other treatment facilities with minimal monitoring or through self-quarantine. High-risk group who are likely to require oxygen therapy can be assigned first to hospitals and receive priority care in early stage. With this scoring system, hospitals can prevent unnecessary overuse of medical care in limited resource settings (e.g., during massive outbreaks), and may reduce mortality through effective allocation of medical resources. During the peak of the Covid-19 outbreak in South Korea, several patient deaths occurred in their homes while waiting to be hospitalized after receiving Covid-19 diagnosis due to shortage of beds. By triaging patients using the KDDH scoring system, such mortality cases could be reduced and be prevented in future outbreaks.

This study has several limitations. First, some study showed that severity of COVID-19 infection is related with smoking, obesity, time between symptoms and hospital admission<sup>15,16</sup>. But there is no data collection of these variables in this study. Second, this study has limitation of being a retrospective, single center cohort study conducted in Daegu, South Korea. However, this hospital treated the largest number of patients with Covid-19, thus its patient data is well-representative of the entire country. Moreover, it is one of the few studies conducted outside of China to develop and validate a risk scoring model for Covid-19 patients. A substantial number of patients were enrolled and followed up for more than 4 weeks without any change observed in final patient outcomes, and the scoring system was well-validated.

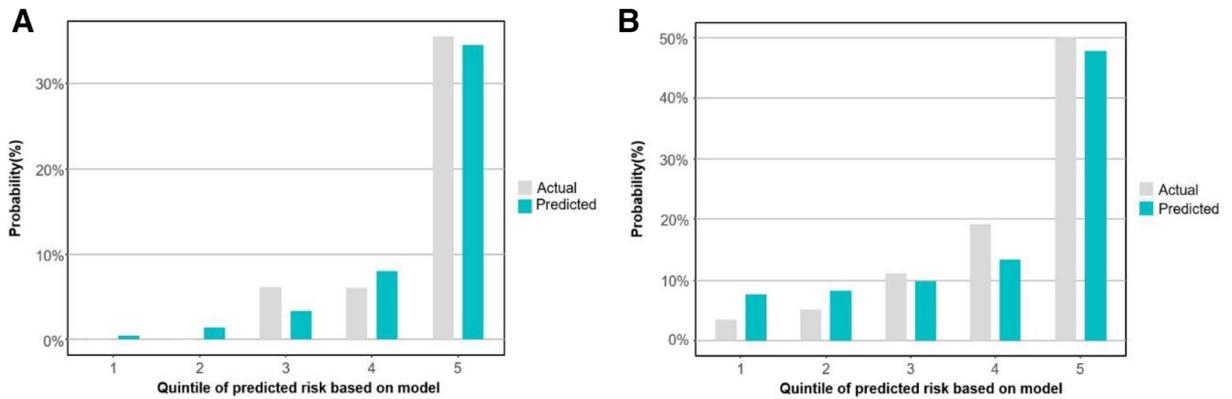
In our experience, based on the KDDH scoring system, the low-risk patients received care in the mild patient ward where oxygen therapy was unavailable. The high-risk patients were admitted in the main ward where oxygen therapy and close monitoring could be provided. By concentrating medical personnel who can provide critical care at the main ward, we could treat the critically ill patients effectively.



**Figure 2.** Discrimination ability of the severe pneumonia in prediction model in patients with Covid-19. Receiver operating characteristics (ROC) curves for predictive value in the (A) development cohort: fitted value, (B) development cohort: KDDH score, (C) validation cohort: fitted value, and (D) validation cohort: KDDH score.

|                                       | Development cohort (n = 421) |                | Validation cohort (n = 140) |               |
|---------------------------------------|------------------------------|----------------|-----------------------------|---------------|
| AUC (95% CI)                          |                              |                |                             |               |
| Fitted value                          | 0.886 (0.836–0.937)          |                | 0.826 (0.730–0.922)         |               |
| KDDH Score                            | 0.884 (0.833–0.934)          |                | 0.828 (0.733–0.923)         |               |
| Cut-off Points 9 (stable/progression) |                              |                |                             |               |
| High-risk (Points ≥ 9)                | 31 TP (27.7%)                | 81 FP (72.3%)  | 17 TP (41.5%)               | 24 FP (58.5%) |
| Low-risk (Points < 9)                 | 6 FN (1.9%)                  | 303 TN (98.1%) | 7 FN (7.1%)                 | 92 TN (92.9%) |
| Sensitivity (%)                       | 83.8%                        |                | 70.8%                       |               |
| Specificity (%)                       | 78.9%                        |                | 79.3%                       |               |
| Positive predictive value (%)         | 27.7%                        |                | 41.5%                       |               |
| Negative predictive value (%)         | 98.1%                        |                | 92.9%                       |               |
| Positive likelihood ratio             | 3.972                        |                | 3.424                       |               |
| Negative likelihood ratio             | 0.206                        |                | 0.368                       |               |

**Table 3.** Performance for early prediction of severe Covid-19. FN, false negative; FP, false positive; TN, true negative; TP, true positive.



**Figure 3.** Calibration ability of the severe pneumonia prediction model in patients with Covid-19. The calibration bar plot in the development cohort (A) and the validation cohort (B).

In summary, the presented KDDH scoring system was validated to be a highly informative and successful risk stratification tool to identify Covid-19 patients at higher risk of progression to severe pneumonia. Early adoption of this scoring system can assist optimal usage of limited medical resources in different health facility settings that are undergoing rapid Covid-19 outbreaks.

### Data availability

The datasets generated and/or analyzed during the present study are available from the corresponding author on reasonable request.

Received: 19 June 2020; Accepted: 29 April 2021

Published online: 30 March 2022

### References

- World Health Organization. *WHO Director-General's Opening Remarks at the Media Briefing on COVID-19—11 March 2020*. (WHO, 2020).
- Wu, Z. & McGoogan, J. M. Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China. *JAMA* **323**, 1239–1242. <https://doi.org/10.1001/jama.2020.2648> (2020).
- Wang, D. *et al.* Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus-infected pneumonia in Wuhan, China. *JAMA* **323**, 1061–1069. <https://doi.org/10.1001/jama.2020.1585> (2020).
- Wu, C. *et al.* Risk factors associated with acute respiratory distress syndrome and death in patients with coronavirus disease 2019 Pneumonia in Wuhan, China. *JAMA Intern. Med* **180**, 934–943. <https://doi.org/10.1001/jamainternmed.2020.0994> (2020).
- Guan, W.-J. *et al.* Clinical characteristics of coronavirus disease 2019 in China. *N. Engl. J. Med.* **382**, 1708–1720. <https://doi.org/10.1056/NEJMoa2002032> (2020).
- 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* **395**, 1054–1062. [https://doi.org/10.1016/S0140-6736\(20\)30566-3](https://doi.org/10.1016/S0140-6736(20)30566-3) (2020).
- Gandhi, R. T., Lynch, J. B. & Rio, C. D. Mild or moderate Covid-19. *N Engl J Med* **383**, 1757–1766. <https://doi.org/10.1056/NEJMc2009249> (2020).
- World Health Organization. *CLINICAL management of Severe Acute Respiratory Infection When Novel Coronavirus (2019-nCoV) Infection is Suspected: Interim Guidance, 28 January 2020* (WHO, 2020).
- Wang, L. *et al.* Coronavirus disease 2019 in elderly patients: Characteristics and prognostic factors based on 4-week follow-up. *J. Infect.* **80**, 639–645. <https://doi.org/10.1016/j.jinf.2020.03.019> (2020).
- Wynants, L. *et al.* Prediction models for diagnosis and prognosis of covid-19 infection: Systematic review and critical appraisal. *BMJ* **369**, m1328. <https://doi.org/10.1101/2020.03.24.20041020> (2020).
- Jiang, X. *et al.* Towards an artificial intelligence framework for data-driven prediction of coronavirus clinical severity. *Comput. Mater. Contin.* **63**, 537–551. <https://doi.org/10.32604/cmc.2020.010691> (2020).
- Ji, D. *et al.* Prediction for progression risk in patients with COVID-19 pneumonia: The CALL score. *Clin. Infect. Dis.* **71**, 1393–1399. <https://doi.org/10.1093/cid/ciaa414> (2020).
- Gong, J. *et al.* A tool to early predict severe 2019-novel coronavirus pneumonia (COVID-19): A multicenter study using the risk nomogram in Wuhan and Guangdong, China. *Clin. Infect. Dis* **71**, 833–840. <https://doi.org/10.1101/2020.03.17.20037515> (2020).
- 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.* **180**, 1081–1089. <https://doi.org/10.1001/jamainternmed.2020.2033> (2020).
- Hu, L. *et al.* Risk factors associated with clinical outcomes in 323 coronavirus disease 2019 (COVID-19) hospitalized patients in Wuhan, China. *Clin. Infect. Dis.* **71**, 2089–2098. <https://doi.org/10.1093/cid/ciaa539> (2020).
- Alaa, A. *et al.* Retrospective cohort study of admission timing and mortality following COVID-19 infection in England. *BMJ Open* **10**, e042712. <https://doi.org/10.1136/bmjopen-2020-042712> (2020).

### Acknowledgements

We are very thankful for all the medical personnel who are fighting against the worldwide menace, Covid-19, people who are sincerely supporting the related medical acts, and people who are continuously following proper instructions to prevent the spread of the infection. Moreover, we thank Keimyung University Dongsan Hospital and Daegu Dongsan Hospital staff for their outstanding devotion to overcome the Covid-19 crisis. Finally, we would like to thank HERINGS for model development and all the statistical analyses.

### Author contributions

J.Y.L., C.H.C. mainly designed and developed the manuscript; B.H.N. conducted statistical analysis and provided input; MK assisted in manuscript development; and J.Y.L., J.H., J.Y.K., M.H. and H.A.K. provided input. All authors reviewed the manuscript.

### Funding

This research was supported by the Bisa Research Grant of Keimyung University in 2015.

### Competing interests

The authors declare no competing interests.

### Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-07610-9>.

**Correspondence** and requests for materials should be addressed to C.-H.C.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022