Original Paper

COVID-19 Outcome Prediction and Monitoring Solution for Military Hospitals in South Korea: Development and Evaluation of an Application

JoonNyung Heo[1*], MD; Ji Ae Park[2*], MS; Deokjae Han[3], MD; Hyung-Jun Kim[3], MD; Daeun Ahn[4], BS; Beomman Ha[1], MD, PhD; Woong Seog[1], MD; Yu Rang Park[2], PhD

[1]Armed Forces Medical Command, Seongnam, Republic of Korea

[2]Department of Biomedical Systems Informatics, Yonsei University College of Medicine, Seoul, Republic of Korea

[3]Department of Internal Medicine, The Armed Forces Capitol Hospital, Seongnam, Republic of Korea

[4]Department of Nursing, The Armed Forces Capitol Hospital, Seongnam, Republic of Korea

*these authors contributed equally

Corresponding Author:
Yu Rang Park, PhD
Department of Biomedical Systems Informatics
Yonsei University College of Medicine
50-1 Yonsei-ro, Seodaemun-gu
Seoul, 03722
Republic of Korea
Phone: 82 2 2228 2493
Email: yurangpark@yuhs.ac

Abstract

Background: COVID-19 has officially been declared as a pandemic, and the spread of the virus is placing sustained demands on public health systems. There are speculations that the COVID-19 mortality differences between regions are due to the disparities in the availability of medical resources. Therefore, the selection of patients for diagnosis and treatment is essential in this situation. Military personnel are especially at risk for infectious diseases; thus, patient selection with an evidence-based prognostic model is critical for them.

Objective: This study aims to assess the usability of a novel platform used in the military hospitals in Korea to gather data and deploy patient selection solutions for COVID-19.

Methods: The platform's structure was developed to provide users with prediction results and to use the data to enhance the prediction models. Two applications were developed: a patient's application and a physician's application. The primary outcome was requiring an oxygen supplement. The outcome prediction model was developed with patients from four centers. A Cox proportional hazards model was developed. The outcome of the model for the patient's application was the length of time from the date of hospitalization to the date of the first oxygen supplement use. The demographic characteristics, past history, patient symptoms, social history, and body temperature were considered as risk factors. A usability study with the Post-Study System Usability Questionnaire (PSSUQ) was conducted on the physician's application on 50 physicians.

Results: The patient's application and physician's application were deployed on the web for wider availability. A total of 246 patients from four centers were used to develop the outcome prediction model. A small percentage (n=18, 7.32%) of the patients needed professional care. The variables included in the developed prediction model were age; body temperature; predisease physical status; history of cardiovascular disease; hypertension; visit to a region with an outbreak; and symptoms of chills, feverishness, dyspnea, and lethargy. The overall C statistic was 0.963 (95% CI 0.936-0.99), and the time-dependent area under the receiver operating characteristic curve ranged from 0.976 at day 3 to 0.979 at day 9. The usability of the physician's application was good, with an overall average of the responses to the PSSUQ being 2.2 (SD 1.1).

Conclusions: The platform introduced in this study enables evidence-based patient selection in an effortless and timely manner, which is critical in the military. With a well-designed user experience and an accurate prediction model, this platform may help save lives and contain the spread of the novel virus, COVID-19.

XSL•FO
RenderX

Introduction

COVID-19, which is caused by the SARS-CoV-2 virus, has officially been declared a pandemic [1]. Despite the effort, there were still casualties caused by the disease. The mortality rate of the disease varies between countries and even between cities in China [2]. There are speculations that these mortality rate discrepancies between regions are due to the availability of medical resources in that area. For instance, the mortality rate difference between the city of Hubei and other provinces is significant (~3% for Hubei and 0.7% for the rest of China), and plotting the cases per population against mortality rates shows a positive correlation [2].

The spread of the virus is currently placing sustained demands on public health systems [3]. At present, there is an imbalance in the supply and demand of medical supplies, and many efforts are being made to solve this problem [4]. Although it is ideal if the increase in medical supplies could meet the high demand, this seems to be difficult for many underdeveloped countries.

Adjusting demand, such as by selecting the patients for diagnosis and treatment, can also be an option considering that the rates for severe cases are relatively low, with severe cases at 14% and critical cases at 5% based on Chinese reports [5]. Comparably, rates of the asymptomatic cases are surprisingly high, with some reporting up to 75% [6-8]. In this context, the Centers for Disease Control and Prevention (CDC) recommends that not all patients be hospitalized, considering the insufficient medical infrastructure. However, there is no definite guideline regarding patient selection due to insufficient evidence on which patients need professional care.

Military personnel are especially at risk, since the viral spread is the most critical in settings where people are living closely in groups [9,10]. The adenovirus infection is one of the examples of an infectious disease that is easily spread in the military. Additionally, soldiers are frequently relocated worldwide, which may accelerate the spread of the disease. For instance, in 2007, basic military trainees were infected with the adenovirus and were sent to multiple sites throughout the world, causing an outbreak in those military bases [9].

The confusion, urea, respiratory rate, blood pressure, and 65 years or older (CURB-65) score is a widely known scoring system for predicting the outcome of patients with pneumonia [11]. The presence of confusion, blood urea nitrogen, respiratory rate, systolic and diastolic blood pressure, and age are assessed to calculate a score ranging from 0 to 6, with 6 indicating the highest severity. The American Thoracic Society guidelines suggest that patients with CURB-65 scores of less than 2 are at a low risk of death and may be managed as outpatients. In response to the outbreak, this score was used for patients with COVID-19 [12]. However, even though patients with higher CURB-65 scores were more likely to be severe, more than 20%

of the patients with lower scores also had poor outcomes, suggesting the need for a new system to guide patient selection.

Previous studies have developed prognostic models to predict mortality risk, hospital stay, and progression to severe state [13-21]. The largest sample size of these studies was 577 patients from China [19]. Ideally, prediction models should be based on a large data set covering multiple nations and races for vast adaptation. Four of these studies presented models (eg, decision tree, nomogram, and scoring rule) that could be used in clinical practice [14,16,17,21].

When a prediction model is being developed and deployed in the real world, multiple factors should be considered. As previously stated, the included samples should be from multiple centers across multiple nations to account for disparities between races. Additionally, the model should provide a practical and irreplaceable value to its users. The objective of the model should be clear and should be able to answer a major question in the clinical field. The model's ease of use should also be considered. However, these requirements are difficult—if not impossible—to meet in a situation where a novel infection is spreading worldwide.

In this study, we present a platform that provides outcome prediction and status monitoring for patients with COVID-19 that is consistently enhanced with data collected based on the use of the model.

Methods

Development of the Platform Structure

The platform structure was designed to primarily serve the users by providing the results of the prediction models and sending the results to the appointed physician. The data sent by the user to acquire the results were collected by the server with the user's consent. The newly collected data was used for training the models to further enhance its predictability and generalizability, thereby completing the virtuous cycle. There are two applications employed to serve this purpose: (1) a patient's application and (2) a physician's application.

The main functionality of the patient's application is to provide outcome prediction results to the patient with their general information, previously diagnosed diseases, symptoms, and body temperature. Specifically, the purpose of the outcome prediction is to guide patient selection for medical resource allocation. Thus, the main target of the prediction model is to determine whether the patient will need professional care.

The application receives the variables that the patient provides to acquire the prediction results, sends them to the central server, and shares them with the appointed physician for remote monitoring. When the patient visits a health care facility, the patient can pair their application with the physician's application to keep the physician notified of the change in the patient's

symptoms without needing any direct contact. The physician's application then receives a registration code that can be created from the patient's application. Upon registering a patient, the

health care worker can receive all the records and prediction results from the patient's application. This process is depicted in Figure 1.

Figure 1. Diagram of the whole structure of the platform.
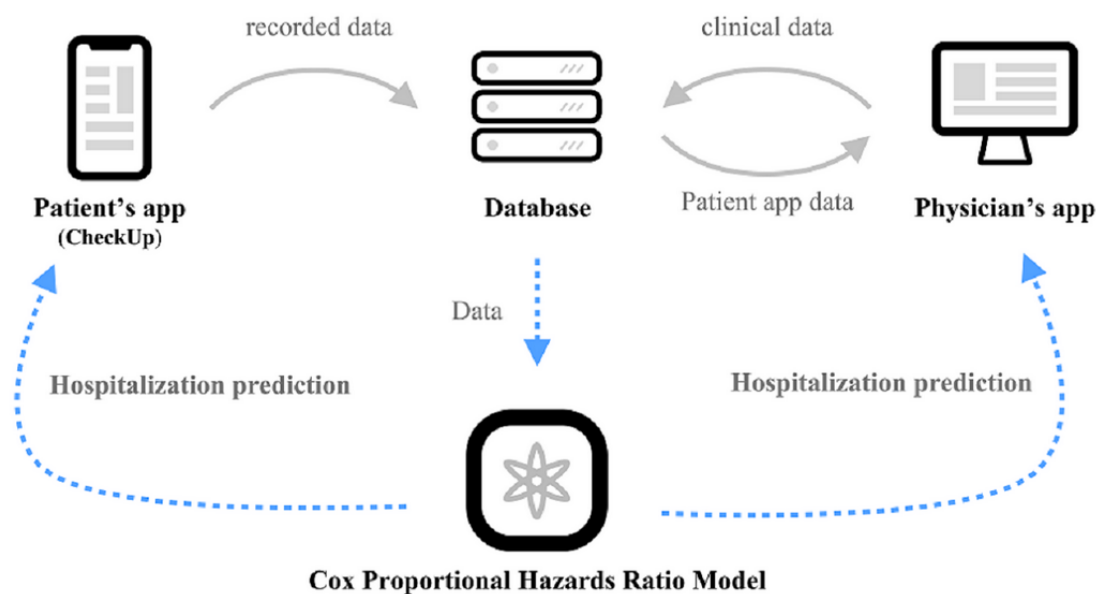
Three physicians reviewed previous publications related to COVID-19 to select the appropriate variables for collection [22-24]. On the basis that this novel virus might have similar features to the pre-existing viruses that cause viral pneumonia, the characteristics of other viral pneumonias were also considered. The selection result was then reviewed by two physicians who are currently directly involved with treating patients with COVID-19 (Multimedia Appendix 1).

Participants and Data Collection for Outcome Prediction

All patients with COVID-19 admitted to four centers (two military hospitals and two civilian hospitals) in Korea were included in the study. A total of 246 patients were included, and data collection ran from February 6, 2020, to April 2, 2020. There were no exclusion criteria.

The candidate risk factors for the patient's application included patient demographic characteristics (age, gender, and BMI), past history (asthma, chemotherapy, chronic kidney disease, hypertension, chronic obstructive pulmonary disease, immunosuppressant use, cardiovascular disease, chronic liver disease, and diabetes), patient symptoms (anosmia, rhinorrhea, chest pain, phlegm, chills, physical status, cough, pneumonia, diarrhea, pneumonia, antipyretic, dyspnea, feverish, headache, muscle pain, nausea or vomiting, tired or lethargy, and sore throat), social history (visit to a region with an outbreak, direct contact with a patient with COVID-19, smoking history, household member confirmed with COVID-19, and household member under self-isolation), and body temperature during hospitalization.

The outcome measured in this study was the use of an oxygen supplement monitored up to April 6, 2020—the final date of follow-up. The length of time (days) from the date of

hospitalization to the date of first use of an oxygen supplement was considered as our target outcome. Oxygen supplement was chosen as the outcome measure since it may be able to represent the minimal treatment required for hospital admission.

Statistical Analysis

Descriptive statistics were obtained for all study variables. Continuous data were expressed as mean (SD) values, and categorical data were expressed as proportions.

A survival curve for overall patients was plotted using the Kaplan-Meier method. Using all candidate risk factors, we conducted a univariate analysis by a Cox proportional hazards model to select factors. The factors with $P<.05$ were included in the final model. The body temperature collected during hospitalization was used as the time-dependent variable in the Cox proportion hazards model [25]. By generating time intervals during the day after the date of hospitalization, the average body temperature recorded at each time interval (day) was calculated. That is, the body temperature was considered as a value that changed every day. In case of missing body temperature, the last observation carried forward approach was used. A multivariate Cox proportional hazard model (with time-dependent variables) using selected factors was considered as the prediction model for the patient's application. The multicollinearity between the factors included in the final model was confirmed through variance inflation factors (VIFs), which quantify the severity of multicollinearity in an ordinary least squares regression analysis. We considered that VIF of 5 or above indicates a multicollinearity problem [26]. Also, additionally, association between factors was examined through phi coefficient between two binary factors, point-biserial correlation between a continuous and a binary factor, and pearson correlation between two continuous factors. To calculate VIF through the ordinary least square regression analysis and

to measure correlation between factors, the median body temperature during hospitalization was considered for body temperature with time-varying characteristics.

We calculated the Harrell C statistic to assess the overall predictive accuracy and time-dependent area under the receiver operating characteristic curve (area under the curve [AUC]) [27] to assess the predictive accuracy over the entire follow-up period. The AUC results of the Cox model, which vary depending on the time point to be evaluated, were summarized by days. We also performed the likelihood ratio, Wald, and score tests to measure overall goodness of fit based on the omnibus test of model coefficients. The chi-square goodness-of-fit test proposed by Schoenfeld [28] was used for the proportional hazards. Since body temperature was considered as a time-varying variable updated every day in the model, we did not check the assumption for body temperature. A random subsampling [29] was used to estimate the internal validity of the final Cox proportional hazards model, with 50 repeated samplings dividing the train and test sets into 2:1. The sampling was performed by stratifying according to the occurrence of the outcome event to establish balance in the test and train sets. An average time-dependent AUC based on the 50 repetitions was used as a result of the validation.

The Cox proportional hazards model was considered as a prediction model for the patient's application, but we additionally constructed a logistic model as an alternative model by using selected factors from the final Cox model. The logistic model was designed to predict oxygen supplement use during hospitalization.

All P values were two-sided, and $P<.05$ was considered statistically significant. Statistical analysis was performed using R 3.6.0 (R Foundation for Statistical Computing) [29].

Evaluation of the Physician's Application

A usability study was conducted to evaluate the usability of the health care worker's application wherein participants had to be medical doctors with more than 1 year of clinical practice. A total of 50 physicians were recruited from a public website, and there was no exclusion or selection of the participants. The participants were introduced to the entire platform and briefed about its objectives. After a thorough explanation of the study, a consent form was signed by the participants. They were then instructed to sign up for the service through the physician's application and then add a prepared sample patient case that was provided. Afterward, participants were required to complete several tasks: review the symptoms that the sample case reported through the patient's application, review the results from the outcome prediction model of the registered patient, and add additional factitious clinical variables for the patient from the variable input form. The result from the model was presented as the probability of requiring an oxygen supplement, expressed as a number between 0%-100%. The participants were educated on the intended clinical utility of the result from the prediction model to provide supportive information that can be used during patient selection. However, the usefulness of the prediction model could not be assessed since the data were not from a real patient. The participants were instructed to fill the factitious values of the patients' clinical variables required by the physician's input form. The study protocol was approved by the institutional review board at the Yonsei University College of Medicine (4-2020-0351).

Following the completion of the tasks, the Post-Study System Usability Questionnaire (PSSUQ) was used to acquire the participants' responses [30]. The results from the PSSUQ were analyzed using statistical representative indexes for each question.

Results

Platform Structure

The platform consists of three parts: (1) the patient's application, (2) the physician's application, and (3) the model application programming interface. Both the patient's and the physician's applications are deployed on the web. The web platform was chosen to be able to enable instant modification upon deployment and maximize the availability on variable devices. Considering the majority of devices the users will be using, the patient's application is configured to be best viewed on a mobile device and the physician's application on a desktop. The diagram of the entire structure is illustrated in Figure 2.

Figure 2. Screenshots showing how patient recorded data is shown in the physician's app to monitor the patient's status.
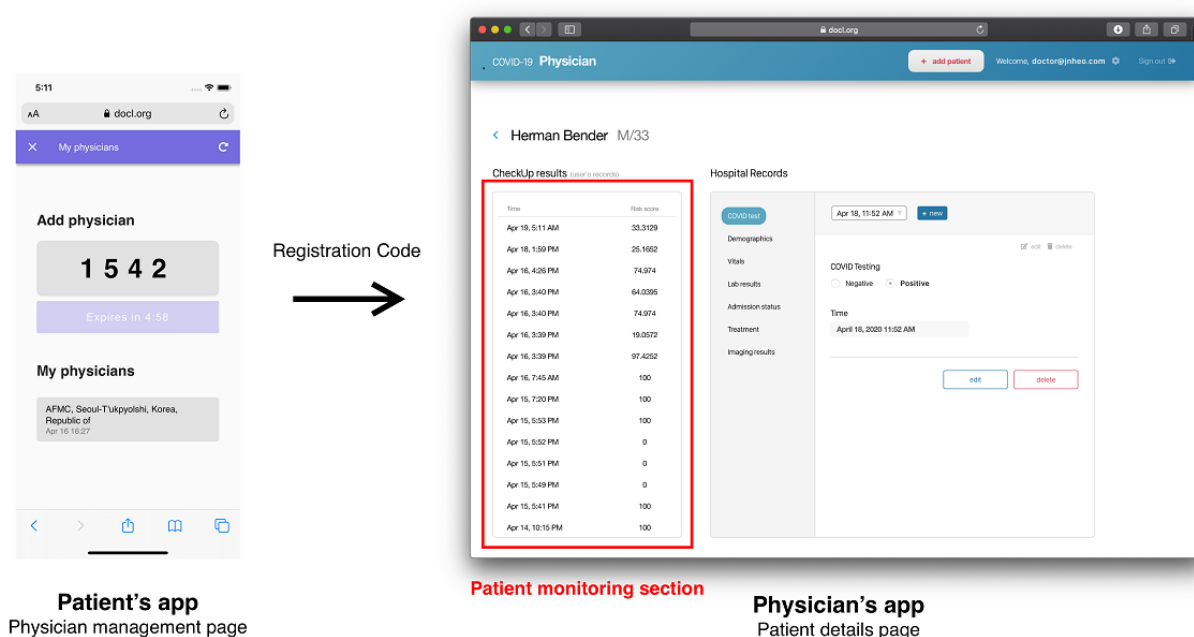
This platform is registered in the World Health Organization Digital Atlas for COVID-19 solutions [31].

Real-World Use of the Platform

A patient is advised to use the patient's application when they are officially diagnosed with COVID-19. They are also advised to use the application every day and provide information about their daily symptoms and body temperature. If the result from the application shows a high likelihood of requiring hospitalization and the patient is not currently admitted to a hospital, then the patient is advised to contact a nearby health care professional. If there is a previously appointed physician monitoring the patient's daily symptoms, then the physician can check the daily status of the patient. Using the results of the application, the physician is then able to make a proper clinical decision about whether the patient will need hospitalization. The result that the physician receives is the raw probability calculated from the prediction model. No cut-off was set, since the cut-off will be affected by environmental factors, such as the availability of hospital beds or the regional number of confirmed patients.

Development of the Outcome Prediction Model

The baseline characteristics of all 246 patients are shown in Multimedia Appendix 2. The mean (SD) of age (year) and BMI (kg/m^2) were 40.72 (SD 17.10) years and 3.21 (SD 3.28) kg/m^2, respectively. A total of 167 (67.89%) patients were male. The mean (SD) of the max and median body temperature during follow-up were 37.32 °C (SD 0.56) and 36.75 °C (SD 0.42), respectively. The most common symptom was coughing (95 patients, 41.13%), and 192 (83.12%) patients visited a region with a COVID-19 outbreak.

One patient with missing date of hospitalization among 246 patients was excluded to estimate the survival rate for oxygen supplement. During 2469.32 person-days of follow-up, oxygen supplement was observed in 18 (7.35%) patients. The median (IQR) time from hospitalization to censoring was 17.64 (IQR 14.24-21.70) days. The Kaplan-Meier estimates of the 3-day and 5-day survival rate were 0.947 (95% CI 0.919-0.975) and 0.930 (95% CI 0.898-0.963), respectively.

Table 1 shows results of the univariate and multivariate Cox proportional hazards models. In the univariate analysis, 10 of the candidate risk factors showed results of $P<.05$ (Table 1). It was confirmed that hypertension (hazard ratio [HR] 3.792, 95% CI 1.423-10.109), cardiovascular disease (HR 12.413, 95% CI 4.069-37.869), predisease physical status (HR 1.854, 95% CI 1.24-2.773), dyspnea (HR 13.498, 95% CI 4.527-40.252), feeling feverish (HR 6.282, 95% CI 2.054-19.213), chills (HR 5.727, 95% CI 1.924-17.048), tiredness or lethargy (HR 6.083, 95% CI 1.989-18.607), older age (HR 1.075, 95% CI 1.052-1.098), and higher body temperature (HR 13.147, 95% CI 6.849-25.237) were risk factors for earlier oxygen supplement, while visiting a region with an outbreak (HR 0.291, 95% CI 0.095-0.89) was confirmed as a risk factor that decreased the risk of oxygen supplement.

Table 1. Hazard ratio of univariate and multivariate Cox proportional hazards model.

Factors	Univariate model, HR[a] (95% CI)	Multivariate model[b]	
		HR (95% CI)	P value
Age (years)	1.075 (1.052-1.098)	1.035 (0.983-1.089)	.47
Body temperature[b] (°C)	13.147 (6.849-25.237)	17.431 (2.856-106.371)	.01
Hypertension (yes)	3.793 (1.423-10.109)	0.562 (0.033-9.437)	.81
CVD[c] (yes)	12.413 (4.069-37.869)	0.217 (0.008-6.111)	.49
Visit to a region of outbreak (yes)	0.291 (0.095-0.89)	3.381 (0.133-86.084)	.28
Physical status	1.854 (1.24-2.773)	4.259 (1.679-10.802)	.007
Dyspnea (yes)	13.498 (4.527-40.252)	3.878 (0.454-33.111)	.43
Feverish (yes)	6.282 (2.054-19.213)	0.321 (0.05-2.073)	.26
Chills (yes)	5.727 (1.924-17.048)	0.905 (0.049-16.705)	.95
Tired/lethargic (yes)	6.083 (1.989-18.607)	1.506 (0.174-13.019)	.62

[a]HR: hazard ratio.

[b]Cox proportional hazards model with time-dependent variable.

[c]CVD: cardiovascular disease.

As a result of multivariate analysis including all 10 factors, body temperature (HR 17.431, 95% CI 2.856-106.371) was found to be the most powerful risk factor for oxygen supplement. Physical status remained statistically significant as a factor (P=.007), while the other eight factors were not statistically significant.

The proportional hazards assumption on the factors included in the multivariate model (except for body temperature, since it is considered as a time-varying variable) was checked. Results showed that all factors satisfied the assumption (Multimedia Appendix 3). The multicollinearity and association between the factors included in the multivariate model were not detected (Multimedia Appendices 4 and 5).

To confirm the effects of body temperature in the multivariate model, we considered a case for a fully active man 40 years of age without any past history or symptoms and compared survival rates according to body temperature (Figure 3). In cases of temperatures of 37 °C, 38 °C, 38.5 °C, and 39 °C, the 5-day survival rates were 0.994 (95% CI 0.9810-1), 0.906 (95% CI 0.7071-1), 0.663 (95% CI 0.1782-1), and 0.180 (95% CI 0.0002-1), respectively. At 39 °C, the 5-day survival rate rapidly decreased.

Figure 3. Survival rate according to body temperature in predictive model present in the patient's application. The 5-day survival rates for each initial body temperature are shown.
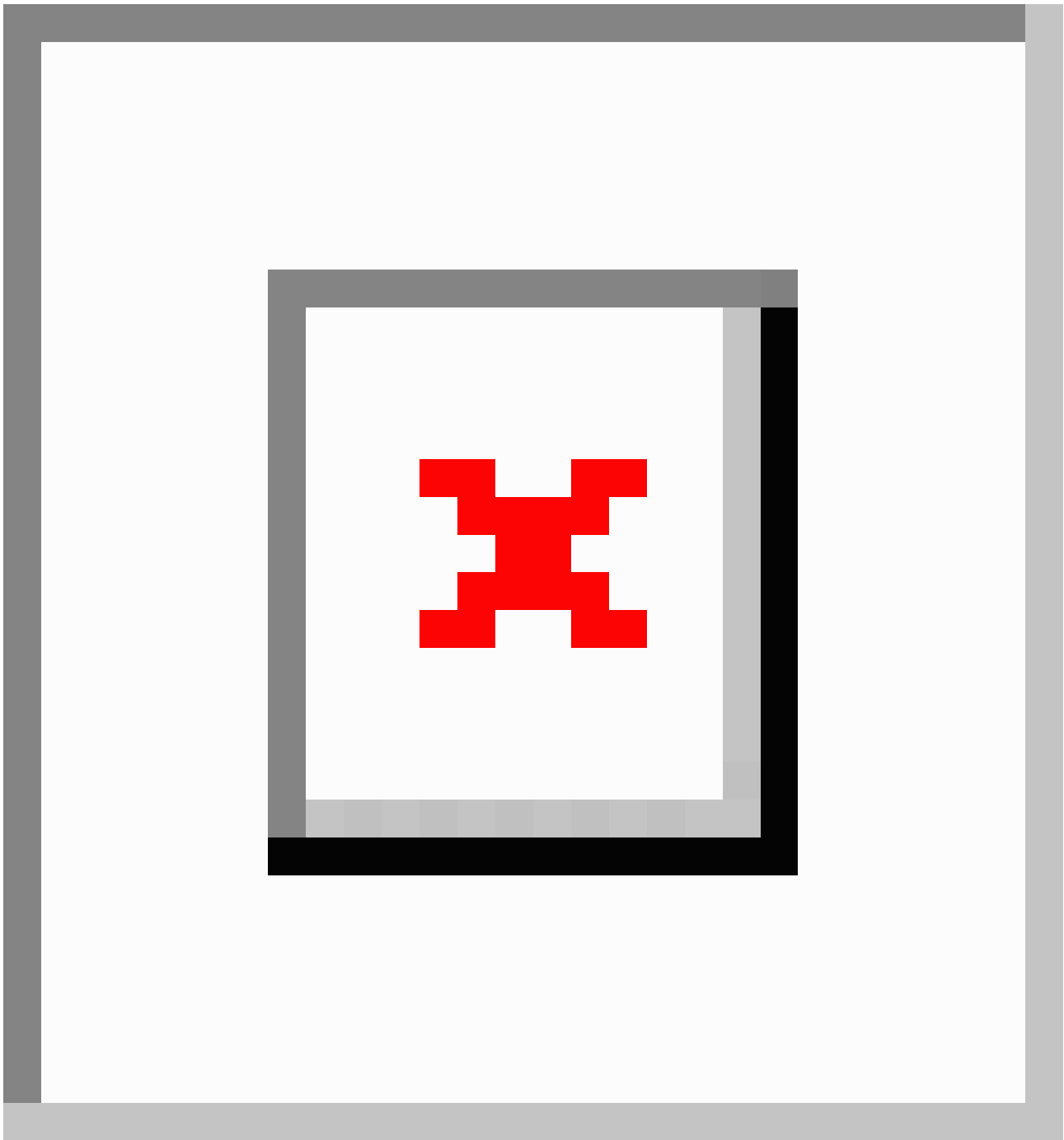
For the multivariate model, the overall C statistic was 0.963 (95% CI 0.936-0.99), and the time-dependent AUC ranged from 0.976 at day 3 to 0.979 at day 9 (Figure 4). Multimedia Appendix 6 shows the time-dependent AUC ranging from day 1 to day 10. The value for each day was above 0.96. The AUC of 8 days or more were the same because the outcome event did not occur after 7.25 days.

Figure 4. Time-dependent area under the receiver operating characteristic curve at 3, 6, and 9 days from the prediction model present in the patient's application.
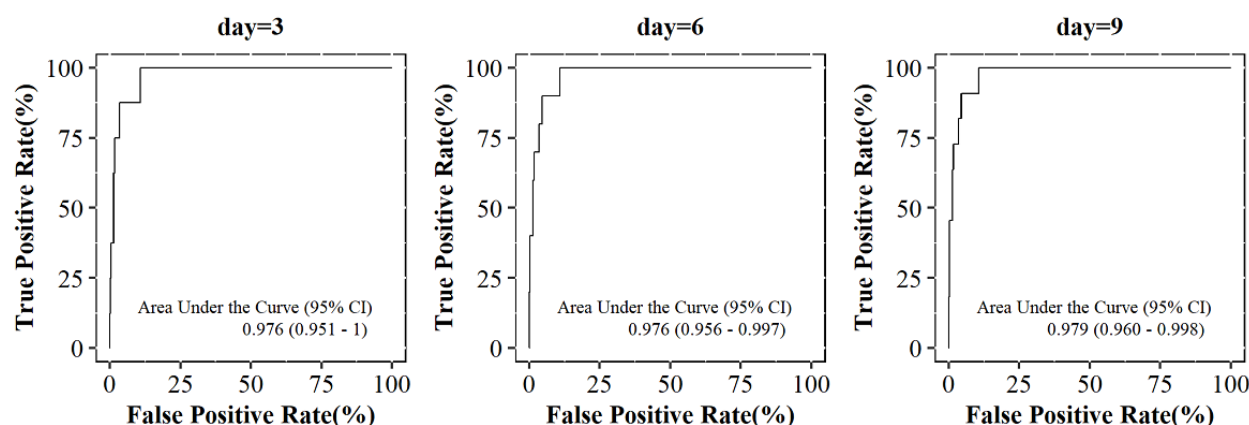
As per the result of internal validation for estimating the performance and validity of the multivariate Cox proportional hazards model in Table 1, the average time-dependent AUC ranging from day 1 to day 10 showed a range from a minimum of 0.749 to a maximum of 0.764 (Multimedia Appendix 7).

The results from the model are displayed at the result section of the application when the user has completed a series of questionnaires. The results page shows the 7day survival result based on the received variables.

As results of the multivariate logistic model, body temperature (odds ratio [OR] 19.106, 95% CI 1.587-229.961), physical status (OR 5.145, 95% CI 1.539-17.205), and age (OR 1.099, 95% CI 1.028-1.174) were statistically associated with oxygen supplementation during hospitalization (Multimedia Appendix 8). The P values of the likelihood ratio and score tests were less than .001, and the Hosmer and Lemeshow test showed a P value

of .99. In addition, the AUC was 0.973 (95% CI 0.9413-1; Multimedia Appendix 9).

Platform Evaluation

The usability study was performed with 50 participants from May 1, 2020, to May 31, 2020. Participants had a mean age of 35.9 years, and 35 (70.0%) were men. All participants were able to complete the given tasks easily. The detailed results are shown in Table 2. The participants were instructed to answer the PSSUQ on a scale of 1-7, with "1" being the most agreeable to the statement. The lower the values of the answers, the better the user experience was. Overall, the participants were satisfied with the app, with the overall average of the responses being 2.2 (SD 1.1). The most disagreeable statement from the questionnaire was the seventh statement ("The system gave error messages that clearly told me how to fix problems"), which had a mean response of 2.7. However, no error was experienced during the study process, thus many of the participants chose "4" as a response, which stood for "neutral."

Table 2. Usability study results (score ranging from 1 to 7, one being "strongly agree").

Statements	Score, mean (SD)
1. Overall, I am satisfied with how easy it is to use this system.	2.0 (1.0)
2. It was simple to use this system.	2.0 (1.1)
3. I was able to complete the tasks and scenarios quickly using this system.	2.2 (1.1)
4. I felt comfortable using this system.	2.2 (1.2)
5. It was easy to learn to use this system.	2.0 (1.1)
6. I believe I could become productive quickly using this system.	2.0 (1.1)
7. The system gave error messages that clearly told me how to fix problems.	2.7 (1.3)
8. Whenever I made a mistake using the system, I could recover easily and quickly.	2.3 (1.2)
9. The information such as online help, on-screen messages, and other documentation provided with this system was clear.	2.0 (1.0)
10. It was easy to find the information I needed.	2.3 (1.2)
11. The information was effective in helping me complete the tasks and scenarios.	2.1 (1.1)
12. The organization of information on the system screens was clear.	2.0 (1.1)
13. The interface of this system was pleasant.	2.1 (1.2)
14. I liked using the interface of this system.	2.1 (1.2)
15. This system has all the functions and capabilities I expect it to have.	2.4 (1.2)
16. Overall, I am satisfied with this system.	2.1 (1.0)

Discussion

Principal Findings

In this study, we developed a platform to provide the users with the most up to date, evidence-based prediction model that can guide them in making decisions on whether to seek professional care or hospitalization, using only the variables provided by the users. The result provided to the users is the probability of the patient requiring oxygen supplement. With this, patients can effectively and accurately monitor themselves on whether they will need hospitalization during self-quarantine. Physicians can use the result from the prediction model for patient selection for hospitalization and risk assessment when needed. The platform collects the data required to calculate the prediction results and uses the data to update the prediction results given directly to the users. The developed Cox model in this study has a high accuracy with an AUC of 0.97 or higher. Our model showed that body temperature was the most important factor for oxygen supplement. A usability study was performed with PSSUQ, which showed that the participants were generally satisfied with the application.

Timely Patient Selection and the Military

The platform is currently used in military hospitals in Korea. Collection of data and deployment of the outcome prediction model is essential in military settings. If a soldier were to be confirmed with the disease, an exceptionally large number of people would be at risk. In contrast, the patients are likely to be considered low risk, that is, having low probability of requiring professional care, since most of them will be younger and healthier than the general population. This special circumstance, where an explosive spread of the virus is expected but only a few will need special care, will pose a greater importance in the selection of patients for both diagnosis and treatment. A large fraction of the confirmed patients will not be hospitalized, so a close and efficient method for monitoring these patients is necessary. However, the platform was developed considering its use outside of the military when needed. Therefore, patients from two public hospitals were included in the study to address this possibility.

The platform is tailored to fit needs in situations like the COVID-19 pandemic, where data of the novel disease is scarce, and the disease is spreading so fast that traditional clinical trials are not timely enough. Clinical trials are carefully designed to minimize bias and clearly prove a hypothesis. Accuracy and reliability are critical in the medical field, but progressive measures may be required during desperate times. Furthermore, real-world data may be more representative of the patient populations in the clinical field [32]. However, the results derived from this platform should be used in a complementary manner, especially if there is more reliable evidence such as a randomized controlled trial.

The main goal of this platform is to provide the physician with a supportive measure to assess the patient's risk of requiring professional care. The probability of the patient requiring admission is provided to the physician in the form of a percentage (ie, 0%-100%). The model does not decide if the patient will need hospitalization but provides the risk information of the patient to the physician so a proper decision can be made. Considering the availability of medical resources and the number of patients confirmed with COVID-19, the physician will have to decide which patient will require hospitalization. The prediction model is expected to help the physician with complementary information for the better prioritization of patients.

Evaluation of the Outcome Prediction Model

The model presented in our study used 246 patients' data from four centers in Korea to predict whether a patient would need hospitalization during the course of the disease. The Cox proportional hazards ratio model was used to account for the time-dependent variable (ie, body temperature), and the model showed a high predictive accuracy (C statistic: 0.963; AUC at 9 days: 0.979). Ten features were selected in the model: age, body temperature, history of hypertension, cardiovascular disease, visit to a region with an outbreak, predisease physical status, dyspnea, feeling feverish, chills, and tiredness or lethargy. Body temperature showed the highest HR (17.431, 95% CI 2.856-106.371), followed by predisease physical status (HR 4.259, 95% CI 1.679-10.802) and dyspnea (HR 3.898, 95% CI 0.454-33.111; Table 1).

The predictors selected for the model are mostly consistent with previous reports. Although age and comorbidities are already well-established risk factors for grave outcomes [33,34], dyspnea was also found to be associated with disease progression [33]. In our study, the effect of body temperature was the most important predictor. For example, a patient who is a fully active man 40 years of age without any past history or symptoms was highly likely to need an oxygen supplement within 5 days with a 5-day survival rate of 0.18 at 39 °C (Figure 3). Abnormal body temperature is a well-known risk factor for grave prognosis in patients with community-acquired pneumonia [35]. Physical status, which is often expressed in the activities of daily living score, is known to be an independent risk factor for mortality among older adult patients with pneumonia [36,37].

When multivariable prediction models are developed, the number of outcome events compared to the number of predictors, referred to as the events per variable (EPV), affects the accuracy of the model. In general, it is known that a reliable sample size is at least 10 events per predictor (variable) in logistic and Cox models [38,39]. In our model, this rule is not satisfied for the 10 predictors included in the final model. The problem of this low EPV was affected by the accuracy and precision of the coefficients. In addition to the sample size, there is also a possibility that the standard error of the coefficient was overestimated due to the multicollinearity [40], and as a result of the VIF, there was no multicollinearity between predictors (Multimedia Appendix 10). Since overfitting most notably occurs when the number of candidate predictors is large relative to the number of outcome events [41], our model could also be an overfitted model. For this reason, the AUC in the total development sample was about 0.97 (Figure 4), but the result in the internal validation was about 0.75 (Multimedia Appendix 7).

Caution should be employed in the interpretation of several risk factors' HR in the developed Cox model. We constructed the multivariate Cox model by selecting only statistically significant variables from the univariate model. In the univariate model, the HRs of all risk factors were calculated in the expected direction for outcome risk. However, in the multivariate model, hypertension, cardiovascular disease, feeling feverish, and visits to a region with an outbreak were calculated in a different direction from the result in the univariate model. As these factors have a relatively weak influence compared to body temperature and physical status, and the statistical significance disappeared, caution should be taken in the interpretation of the estimated HRs in contrast to the univariate model.

In the developed Cox proportional hazards model, it is accurate to define the onset of follow-up as the date when the patient was infected with COVID-19. However, it was difficult to accurately estimate the date of infection. Therefore, the onset of follow-up was defined as the date of hospitalization or the date on which the patient first recorded data in the application. The inability to accurately estimate the period between the date when the patient was infected with COVID-19 and the date considered as the onset of follow-up can be considered as a limitation of the model.

Usability of the Physician's Application

Usability evaluation was performed for the physician's application since the use of the application should be effortless to lessen their overwhelming workload due to the COVID-19 pandemic. Additionally, the application will be used in the setting where a physician must monitor a large number of patients; therefore, the usability will have a major impact. The application is simple, presenting a list of associated patients on the first page and showing details of the patient when clicking on the list. Consequently, the use questionnaire showed promising results, with a mean score of the total PSSUQ being 2.2 (SD 1.1). The time taken to complete the given tasks was not recorded, but many participants had completed the tasks under 10 minutes and were surprised by how simple the tasks were. The most disagreeable question was the seventh question, which inquired whether the error message was easy to understand for the user. Since no participant experienced an error, the result from the question was unreliable.

Other Applications and Outcome Prediction Models for COVID-19

Although not evidence-based, there are some applications that help patients by providing behavioral guidance when they suspect COVID-19 infection. Our initially developed applications, CheckUp Classic and Triage Classic, provided expert opinion-based guidelines in patient selection for testing and patient triage, respectively. These applications received more than 240,000 visits worldwide in a month. Additionally, Apple Inc published a COVID-19 Screening Tool in cooperation with the CDC [42]. A self-triage and self-scheduling tool were also developed based on a well-designed algorithm. This tool is designed to be tethered to the electronic medical records [43].

A few models have been developed to predict the outcome for patients with COVID-19, and a systemic review has been performed [44]. In the review, the authors conclude that all of these models are not useful in the clinical setting. This is primarily due to poor adherence to guidelines and small data size, both of which are unavoidable in a situation where there is not enough time. In a recent editorial, the author emphasizes the importance of sharing data to overcome the limitations of building such a prediction model [45]. However, motive for the researchers to share the hard-earned patient data is weak. Data collection is an additional burden for the already occupied health

XSL•FO

RenderX

care providers. There are a few efforts to collect nationwide or worldwide data for patients with COVID-19 [46,47].

COVID-19 Estimated Risk (COVER) was one of the models developed using the common data model [48]. The model comprises nine variables: age, sex, history of cancer, chronic obstructive pulmonary disease, diabetes, heart disease, hypertension, hyperlipidemia, and kidney disease. It was developed based on 6,869,127 patients with influenza or flu-like symptoms. Although the score itself was not originally developed for patients with COVID-19, it was validated with 43,061 patients with COVID-19. The performance, measured with the AUC, was 0.73-0.81 for COVER-H (predicting hospitalization), 0.73-0.91 for COVER-I (predicting intensive care), and 0.82-0.90 for COVER-F (predicting fatality).

Future Research

Regardless of the accuracy or the representativeness of the data set used for the development of the prediction model, the clinical utility of the platform should be prospectively evaluated. Whether additional information from the prediction model is helpful for the physician or if it is redundant should be further studied. Considering that outcome may vary according to environmental factors, it is questionable if a single prediction model could be generalized worldwide or if there should be separate models for each environment. Currently the platform is not considered for evaluation as an authorized medical device. Extensive prospective research following collection of large data sets to fully represent the target population should be performed.

Limitations

There are limitations in this study. First, the design of the study does not conform to any previous guidelines for clinical studies.

This is due to the difference of scope and objective between traditional clinical studies and this platform. Even so, this study is prone to many types of bias, so the results from this study should be handled with care. Auditing of data is not possible since there is no direct contact with the users, so the completeness or accuracy of the data cannot be ensured. The platform itself is unique, so the study design has never been validated. Considering that the guideline for patient hospitalization is dependent on multiple factors including hospital capacity, governmental guidelines, and patient risk, a single outcome measure cannot be used. Thus, the primary outcome measure—oxygen supplement—is not a definite marker for hospital admission. However, for our data set, all cases were reviewed, and there was no patient that needed hospitalization without receiving oxygen supplement.

Our prediction model was developed from a multicenter data set. However, the model was constructed based on a small data set of a single ethnicity, which can cause selection bias. In addition, since the model has not been built with sufficient data size, there may be a problem in terms of the model's accuracy. With more data acquisition, we will reinforce our model to expand the population applied to the model and increase the generalizable possibilities by performing external validation for other patients.

Conclusion

The platform introduced in this study provides evidence-based decision support to guide patient selection in an effortless and timely manner, which is critical in the military. With a well-designed user experience and an accurate prediction model, this platform may help save lives and contain the spread of the novel virus, COVID-19.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Variables collected for the patient's application.
[DOCX File , 18 KB-Multimedia Appendix 1]

Multimedia Appendix 2

Overall characteristics of patients at admission.
[DOCX File , 21 KB-Multimedia Appendix 2]

Multimedia Appendix 3

Chi-square test by using the Schoenfeld residuals for the proportional hazards.
[DOCX File , 15 KB-Multimedia Appendix 3]

XSL•FO
RenderX

Multimedia Appendix 4

Testing Global Null Hypothesis: Beta=0 of multivariate Cox proportional hazards model.

[DOCX File , 14 KB-Multimedia Appendix 4]

Multimedia Appendix 5

Correlation matrix between 10 predictor variables included in multivariate Cox proportional hazards model.

[DOCX File , 16 KB-Multimedia Appendix 5]

Multimedia Appendix 6

Time-dependent area under the receiver operating characteristic curve ranged from 1-day to 10-day of multivariate Cox proportional hazards model.

[DOCX File , 15 KB-Multimedia Appendix 6]

Multimedia Appendix 7

Average time-dependent area under the receiver operating characteristic curve ranged from 1-day to 10-day using 50 repeated random subsampling to estimate the internal validity.

[DOCX File , 14 KB-Multimedia Appendix 7]

Multimedia Appendix 8

Odds ratio of multivariate logistic model.

[DOCX File , 16 KB-Multimedia Appendix 8]

Multimedia Appendix 9

Results of overall model evaluation, goodness of fit, and predictive accuracy from multivariate logistic model.

[DOCX File , 16 KB-Multimedia Appendix 9]

Multimedia Appendix 10

Results of multicollinearity test with variance inflation factor.

[DOCX File , 15 KB-Multimedia Appendix 10]

References

1.   WHO Timeline - COVID-19. World Health Organization. 2020 Apr 27. URL: https://www.who.int/news-room/detail/27-04-2020-who-timeline---covid-19 [accessed 2020-05-03]

2.   Ji Y, Ma Z, Peppelenbosch MP, Pan Q. Potential association between COVID-19 mortality and health-care resource availability. Lancet Glob Health 2020 Apr;8(4):e480 [FREE Full text] [doi: 10.1016/S2214-109X(20)30068-1] [Medline: 32109372]

3.   Wang C, Horby PW, Hayden FG, Gao GF. A novel coronavirus outbreak of global health concern. Lancet 2020 Feb;395(10223):470-473. [doi: 10.1016/s0140-6736(20)30185-9]

4.   Ranney ML, Griffeth V, Jha AK. Critical supply shortages — the need for ventilators and personal protective equipment during the COVID-19 pandemic. N Engl J Med 2020 Apr 30;382(18):e41. [doi: 10.1056/nejmp2006141]

5.   Wu Z, McGoogan JM. Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China: summary of a report of 72 314 cases from the Chinese Center for Disease Control and Prevention. JAMA 2020 Apr 07;323(13):1239-1242. [doi: 10.1001/jama.2020.2648] [Medline: 32091533]

6.   Day M. Covid-19: identifying and isolating asymptomatic people helped eliminate virus in Italian village. BMJ 2020 Mar 23;368:m1165. [doi: 10.1136/bmj.m1165] [Medline: 32205334]

7.   Wang Y, Liu Y, Liu L, Wang X, Luo N, Li L. Clinical outcomes in 55 patients with severe acute respiratory syndrome coronavirus 2 who were asymptomatic at hospital admission in Shenzhen, China. J Infect Dis 2020 May 11;221(11):1770-1774 [FREE Full text] [doi: 10.1093/infdis/jiaa119] [Medline: 32179910]

8.   Hu Z, Song C, Xu C, Jin G, Chen Y, Xu X, et al. Clinical characteristics of 24 asymptomatic infections with COVID-19 screened among close contacts in Nanjing, China. Sci China Life Sci 2020 May;63(5):706-711 [FREE Full text] [doi: 10.1007/s11427-020-1661-4] [Medline: 32146694]

9.   Trei JS, Johns NM, Garner JL, Noel LB, Ortman BV, Ensz KL, et al. Spread of adenovirus to geographically dispersed military installations, May-October 2007. Emerg Infect Dis 2010 May;16(5):769-775. [doi: 10.3201/eid1605.091633] [Medline: 20409365]

10. Kak V. Infections in confined spaces: cruise ships, military barracks, and college dormitories. Infect Dis Clin North Am 2007 Sep;21(3):773-84, ix [FREE Full text] [doi: 10.1016/j.idc.2007.06.004] [Medline: 17826623]

11. Lim WS, van der Eerden MM, Laing R, Boersma WG, Karalus N, Town GI, et al. Defining community acquired pneumonia severity on presentation to hospital: an international derivation and validation study. Thorax 2003 May;58(5):377-382. [doi: 10.1136/thorax.58.5.377] [Medline: 12728155]

12. Nguyen Y, Corre F, Honsel V, Curac S, Zarrouk V, Fantin B, et al. Applicability of the CURB-65 pneumonia severity score for outpatient treatment of COVID-19. J Infect 2020 Sep;81(3):e96-e98 [FREE Full text] [doi: 10.1016/j.jinf.2020.05.049] [Medline: 32474039]

13. Shi Y, Yu X, Zhao H, Wang H, Zhao R, Sheng J. Host susceptibility to severe COVID-19 and establishment of a host risk score: findings of 487 cases outside Wuhan. Crit Care 2020 Mar 18;24(1):108 [FREE Full text] [doi: 10.1186/s13054-020-2833-7] [Medline: 32188484]

14. Gong J, Ou J, Qiu X, Jie Y, Chen Y, Yuan L, et al. A tool for early prediction of severe coronavirus disease 2019 (COVID-19): a multicenter study using the risk nomogram in Wuhan and Guangdong, China. Clin Infect Dis 2020 Jul 28;71(15):833-840 [FREE Full text] [doi: 10.1093/cid/ciaa443] [Medline: 32296824]

15. Bai X, Fang C, Zhou Y, Bai S, Liu Z, Xia L, et al. Predicting COVID-19 malignant progression with AI techniques. SSRN J Preprint posted online on March 31, 2020. [doi: 10.2139/ssrn.3557984]

16. Yuan M, Yin W, Tao Z, Tan W, Hu Y. Association of radiologic findings with mortality of patients infected with 2019 novel coronavirus in Wuhan, China. PLoS One 2020;15(3):e0230548 [FREE Full text] [doi: 10.1371/journal.pone.0230548] [Medline: 32191764]

17. Yan L, Zhang HT, Goncalves J, Xiao Y, Wang M, Guo Y, et al. Prediction of criticality in patients with severe COVID-19 infection using three clinical features: a machine learning-based prognostic model with clinical data in Wuhan. medRxiv Preprint posted online March 3, 2020. [doi: 10.1101/2020.02.27.20028027]

18. Qi X, Jiang Z, Yu Q, Shao C, Zhang H, Yue H, et al. Machine learning-based CT radiomics model for predicting hospital stay in patients with pneumonia associated with SARS-CoV-2 infection: a multicenter study. medRxiv Preprint posted online March 3, 2020. [doi: 10.1101/2020.02.29.20029603]

19. Lu J, Hu S, Fan R, Liu Z, Yin X, Wang Q, et al. ACP risk grade: a simple mortality index for patients with confirmed or suspected severe acute respiratory syndrome coronavirus 2 disease (COVID-19) during the early stage of outbreak in Wuhan, China. SSRN J Preprint posted online February 28, 2020. [doi: 10.2139/ssrn.3543603]

20. Caramelo F, Ferreira N, Oliveiros B. Estimation of risk factors for COVID-19 mortality - preliminary results. medRxiv Preprint posted online February 25, 2020. [doi: 10.1101/2020.02.24.20027268]

21. Xie J, Hungerford D, Chen H, Abrams ST, Li S, Wang G, et al. Development and external validation of a prognostic multivariable model on admission for hospitalized patients with COVID-19. SSRN J Preprint posted online April 6, 2020. [doi: 10.2139/ssrn.3562456]

22. Zhou F, Yu T, Du R, Fan G, Liu Y, Liu Z, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. Lancet 2020 Mar;395(10229):1054-1062. [doi: 10.1016/S0140-6736(20)30566-3]

23. Li L, Huang T, Wang Y, Wang Z, Liang Y, Huang T, et al. COVID-19 patients' clinical characteristics, discharge rate, and fatality rate of meta-analysis. J Med Virol 2020 Jun;92(6):577-583 [FREE Full text] [doi: 10.1002/jmv.25757] [Medline: 32162702]

24. Jiang F, Deng L, Zhang L, Cai Y, Cheung CW, Xia Z. Review of the clinical characteristics of coronavirus disease 2019 (COVID-19). J Gen Intern Med 2020 May;35(5):1545-1549 [FREE Full text] [doi: 10.1007/s11606-020-05762-w] [Medline: 32133578]

25. Thomas L, Reyes EM. Tutorial: survival estimation for Cox regression models with time-varying coefficients using SAS and R. J Stat Soft 2014;61(1):1-23. [doi: 10.18637/jss.v061.c01]

26. O'brien RM. A caution regarding rules of thumb for variance inflation factors. Qual Quant 2007 Mar 13;41(5):673-690. [doi: 10.1007/s11135-006-9018-6]

27. Heagerty P, Lumley T, Pepe M. Time-dependent ROC curves for censored survival data and a diagnostic marker. Biometrics 2000 Jun;56(2):337-344. [doi: 10.1111/j.0006-341x.2000.00337.x] [Medline: 10877287]

28. Schoenfeld D. Partial residuals for the proportional hazards regression model. Biometrika 1982;69(1):239-241. [doi: 10.1093/biomet/69.1.239]

29. R: A Language and Environment for Statistical Computing. R Core Team. 2020. URL: https://www.r-project.org/

30. Lewis JR. IBM computer usability satisfaction questionnaires: psychometric evaluation and instructions for use. Int J Hum-Comput Interac 1995 Jan;7(1):57-78. [doi: 10.1080/10447319509526110]

31. DOCL outcome prediction platform (doctors on the cloud). Digital Health Atlas. URL: https://digitalatlas.who.int/p/00lNbWygey [accessed 2020-05-03]

32. Blonde L, Khunti K, Harris SB, Meizinger C, Skolnik NS. Interpretation and impact of real-world clinical data for the practicing clinician. Adv Ther 2018 Nov;35(11):1763-1774 [FREE Full text] [doi: 10.1007/s12325-018-0805-y] [Medline: 30357570]

XSL•FO
RenderX

33. Zheng Z, Peng F, Xu B, Zhao J, Liu H, Peng J, et al. J Infect 2020 Aug;81(2):e16-e25 [FREE Full text] [doi: 10.1016/j.jinf.2020.04.021] [Medline: 32335169]

34. Bhargava A, Fukushima EA, Levine M, Zhao W, Tanveer F, Szpunar SM, et al. Predictors for severe COVID-19 infection. Clin Infect Dis 2020 May 30:ciaa674 [FREE Full text] [doi: 10.1093/cid/ciaa674] [Medline: 32472676]

35. Fine MJ, Auble TE, Yealy DM, Hanusa BH, Weissfeld LA, Singer DE, et al. A prediction rule to identify low-risk patients with community-acquired pneumonia. N Engl J Med 1997 Jan 23;336(4):243-250. [doi: 10.1056/NEJM199701233360402] [Medline: 8995086]

36. Torres OH, Muñoz J, Ruiz D, Ris J, Gich I, Coma E, et al. Outcome predictors of pneumonia in elderly patients: importance of functional assessment. J Am Geriatr Soc 2004 Oct;52(10):1603-1609. [doi: 10.1111/j.1532-5415.2004.52492.x] [Medline: 15450034]

37. Mody L, Sun R, Bradley SF. Assessment of pneumonia in older adults: effect of functional status. J Am Geriatr Soc 2006 Jul;54(7):1062-1067 [FREE Full text] [doi: 10.1111/j.1532-5415.2006.00797.x] [Medline: 16866676]

38. Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR. A simulation study of the number of events per variable in logistic regression analysis. J Clin Epidemiol 1996 Dec;49(12):1373-1379. [doi: 10.1016/S0895-4356(96)00236-3]

39. Peduzzi P, Concato J, Feinstein AR, Holford TR. Importance of events per independent variable in proportional hazards regression analysis II. Accuracy and precision of regression estimates. J Clin Epidemiol 1995 Dec;48(12):1503-1510. [doi: 10.1016/0895-4356(95)00048-8]

40. Harrell FE. Regression Modeling Strategies With Applications to Linear Models, Logistic Regression, and Survival Analysis. New York, NY: Springer; 2001.

41. Riley RD, Snell KI, Ensor J, Burke DL, Harrell FE, Moons KG, et al. Minimum sample size for developing a multivariable prediction model: PART II - binary and time-to-event outcomes. Stat Med 2019 Mar 30;38(7):1276-1296 [FREE Full text] [doi: 10.1002/sim.7992] [Medline: 30357870]

42. COVID-19 screening tool. Apple and CDC. URL: https://www.apple.com/covid19/ [accessed 2020-05-04]

43. Judson T, Odisho A, Neinstein A, Chao J, Williams A, Miller C, et al. Rapid design and implementation of an integrated patient self-triage and self-scheduling tool for COVID-19. J Am Med Inform Assoc 2020 Jun 01;27(6):860-866 [FREE Full text] [doi: 10.1093/jamia/ocaa051] [Medline: 32267928]

44. Wynants L, Van Calster B, Collins GS, Riley RD, Heinze G, Schuit E, et al. Prediction models for diagnosis and prognosis of covid-19 infection: systematic review and critical appraisal. BMJ 2020 Apr 07;369:m1328. [doi: 10.1136/bmj.m1328] [Medline: 32265220]

45. Sperrin M, Grant SW, Peek N. Prediction models for diagnosis and prognosis in Covid-19. BMJ 2020 Apr 14;369:m1464. [doi: 10.1136/bmj.m1464] [Medline: 32291266]

46. Kim ES, Chin BS, Kang CK, Kim NJ, Kang YM, Choi J, Korea National Committee for Clinical Management of COVID-19. Clinical course and outcomes of patients with severe acute respiratory syndrome coronavirus 2 infection: a preliminary report of the first 28 patients from the Korean cohort study on COVID-19. J Korean Med Sci 2020 Apr 06;35(13):e142 [FREE Full text] [doi: 10.3346/jkms.2020.35.e142] [Medline: 32242348]

47. Lean Open Survey on SARS-CoV-2 Infected Patients. 2020. URL: https://leoss.net/ [accessed 2020-05-02]

48. Williams RD, Markus AF, Yang C, Salles TD, Duvall SL, Falconer T, et al. Seek COVER: development and validation of a personalized risk calculator for COVID-19 outcomes in an international network. medRxiv Preprint posted online January 1, 2020. [doi: 10.1101/2020.05.26.20112649]

Abbreviations

AUC: area under the curve
CDC: Centers for Disease Control and Prevention
COVER: COVID-19 Estimated Risk
CURB-65: confusion, urea, respiratory rate, blood pressure, and 65 years or older
EPV: events per variable
HR: hazard ratio
OR: odds ratio
PSSUQ: Post-Study System Usability Questionnaire
VIF: variance inflation factor

XSL•FO
RenderX