



OPEN Predicting coronavirus disease 2019 severity using explainable artificial intelligence techniques

Takuya Ozawa^{1,30}, Shotaro Chubachi^{1,30}✉, Ho Namkoong^{2,30}✉, Shota Nemoto³, Ryo Ikegami³, Takanori Asakura^{1,4,5}, Hiromu Tanaka¹, Ho Lee¹, Takahiro Fukushima¹, Shuhei Azekawa¹, Shiro Otake¹, Kensuke Nakagawara¹, Mayuko Watase¹, Katsunori Masaki¹, Hirofumi Kamata¹, Norihiro Harada⁶, Tetsuya Ueda⁷, Soichiro Ueda⁸, Takashi Ishiguro⁹, Ken Arimura¹⁰, Fukuki Saito¹¹, Takashi Yoshiyama¹², Yasushi Nakano¹³, Yoshikazu Muto¹⁴, Yusuke Suzuki^{4,5}, Ryuya Edahiro¹⁵, Koji Murakami¹⁶, Yasunori Sato¹⁷, Yukinori Okada^{15,18,19}, Ryuji Koike²⁰, Makoto Ishii^{1,21}, Naoki Hasegawa², Yuko Kitagawa²², Katsushi Tokunaga²³, Akinori Kimura²⁴, Satoru Miyano²⁵, Seishi Ogawa^{26,27}, Takanori Kanai²⁸, Koichi Fukunaga¹ & Seiya Imoto^{29,30}✉

Predictive models for determining coronavirus disease 2019 (COVID-19) severity have been established; however, the complexity of the interactions among factors limits the use of conventional statistical methods. This study aimed to establish a simple and accurate predictive model for COVID-19 severity using an explainable machine learning approach. A total of 3,301 patients ≥ 18 years diagnosed with COVID-19 between February 2020 and October 2022 were included. The discovery cohort comprised patients whose disease onset fell before October 1, 2020 ($N=1,023$), and the validation cohort comprised the remaining patients ($N=2,278$). Pointwise linear and logistic regression models were used to extract 41 features. Reinforcement learning was used to generate a simple model with high predictive accuracy. The primary evaluation was the area under the receiver operating characteristic curve (AUC). The predictive model achieved an AUC of ≥ 0.905 using four features: serum albumin levels, lactate dehydrogenase levels, age, and neutrophil count. The highest AUC value was 0.906 (sensitivity, 0.842; specificity, 0.811) in the discovery cohort and 0.861 (sensitivity, 0.804; specificity, 0.675) in the validation cohort. Simple and well-structured predictive models were established, which may aid in patient management and the selection of therapeutic interventions.

Keywords COVID-19, Machine learning, Artificial intelligence

¹Division of Pulmonary Medicine, Department of Internal Medicine, Keio University School of Medicine, Tokyo, Japan. ²Department of Infectious Diseases, Keio University School of Medicine, 35 Shinanomachi, Shinjuku-ku, Tokyo 160-8582, Japan. ³Industrial and Digital Business Unit, Hitachi, Ltd, Tokyo, Japan. ⁴Department of Clinical Medicine (Laboratory of Bioregulatory Medicine), Kitasato University School of Pharmacy, Tokyo, Japan. ⁵Department of Respiratory Medicine, Kitasato University, Kitasato Institute Hospital, Tokyo, Japan. ⁶Department of Respiratory Medicine, Faculty of Medicine, Graduate School of Medicine, Juntendo University, Tokyo, Japan. ⁷Department of Respiratory Medicine, Osaka Saiseikai Nakatsu Hospital, Osaka, Japan. ⁸JCHO (Japan Community Health Care Organization, Internal Medicine, Saitama Medical Center, Saitama, Japan. ⁹Department of Respiratory Medicine, Saitama Cardiovascular and Respiratory Center, Saitama, Japan. ¹⁰Department of Respiratory Medicine, Tokyo Women's Medical University, Tokyo, Japan. ¹¹Department of Emergency and Critical Care Medicine, Kansai Medical University General Medical Center, Osaka, Japan. ¹²Respiratory Disease Center, Fukujuji Hospital, Tokyo, Japan. ¹³Department of Internal Medicine, Kawasaki Municipal Ida Hospital, Kawasaki, Kanagawa, Japan. ¹⁴Department of Infectious Diseases, Tosei General Hospital, Aichi, Japan. ¹⁵Department of Statistical Genetics, Osaka University Graduate School of Medicine, Osaka, Japan. ¹⁶Department of Respiratory Medicine, Tohoku University Graduate School of Medicine, Miyagi, Japan. ¹⁷Biostatistics Unit, Clinical and Translational Research Center, Keio University Hospital, Tokyo, Japan. ¹⁸Department of Genome Informatics, Graduate School of Medicine, the University of Tokyo, Tokyo, Japan. ¹⁹Laboratory for Systems Genetics, RIKEN Center for Integrative Medical Sciences, Yokohama, Kanagawa, Japan. ²⁰Health Science Research and Development Center, Tokyo Medical and Dental University, Tokyo, Japan. ²¹Department of Respiratory Medicine, Nagoya University Graduate School of Medicine, Aichi, Japan. ²²Department of Surgery, Keio University School of Medicine, Tokyo, Japan. ²³Genome Medical Science Project (Toyama), National Center for Global Health and Medicine, Tokyo, Japan. ²⁴Institute of

Research, Tokyo Medical and Dental University, Tokyo, Japan. ²⁵M&D Data Science Center, Tokyo Medical and Dental University, Tokyo, Japan. ²⁶Department of Pathology and Tumor Biology, Kyoto University, Kyoto, Japan. ²⁷Institute for the Advanced Study of Human Biology (WPI-ASHBi), Kyoto University, Kyoto, Japan. ²⁸Division of Gastroenterology and Hepatology, Department of Medicine, Keio University School of Medicine, Tokyo, Japan. ²⁹Division of Health Medical Intelligence, Human Genome Center, the Institute of Medical Science, the University of Tokyo, 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-0071, Japan. ³⁰These authors contributed equally: Takuya Ozawa, Shotaro Chubachi, Ho Namkoong and Seiya Imoto. ✉email: bachibachi472000@z6.keio.jp; hounamugun@gmail.com; imoto@hgc.jp

Coronavirus disease 2019 (COVID-19), caused by severe acute respiratory coronavirus 2 (SARS-CoV-2), emerged as a global pandemic in February 2020¹. The clinical course of COVID-19 ranges from mild to severe². A significant number of patients continue to develop severe disease despite the advances in vaccines^{3,4} and therapies^{5,6}; however, most cases are relatively mild. Therefore, identifying patients at risk of developing severe disease who require additional medical resources is crucial.

Various patient-related factors are associated with severe COVID-19, including obesity⁷, advanced age, diabetes mellitus⁸, hyperuricemia⁹, and Krebs von den Lungen-6 (KL-6) levels¹⁰. White blood cell (WBC) count, C-reactive protein (CRP) levels, blood urea nitrogen (BUN) levels, and lactate dehydrogenase (LDH) levels were higher in patients with severe COVID-19^{11–15}. However, relying on a single biomarker may not be sufficient to predict severe disease because these biomarkers reflect different pathological conditions. The usefulness of combinations of biomarkers such as the LDH/albumin (Alb) ratio or the neutrophil/lymphocyte ratio (NLR) has been reported^{16–18}. Nevertheless, determining the optimal combination of biomarkers and patient background factors to create accurate predictive models for COVID-19 severity remains a challenge. Furthermore, for practical use in clinical settings, it is essential that the model can be easily applied through composite indices such as estimated glomerular filtration rate (eGFR)¹⁹ or the Fibrosis-4 (FIB-4) index²⁰ which combine multiple factors.

Predictive models have been established using conventional statistical methods, such as linear and logistic regression^{21–23}; however, these approaches are limited to predicting the outcomes of complex interactions involving numerous factors. Machine learning (ML), a subset of artificial intelligence (AI) that can predict complex and nonlinear outcomes more effectively than classical analysis, has been used to create predictive models of COVID-19 severity^{24,25}.

ML represents a significant innovation in data science that can provide high predictive accuracy; however, the underlying calculations are often unclear, making its translation into clinical practice difficult. Overfitting is another significant challenge associated with ML; a few models show high performance in the initial training data but lose performance during validation²⁶. Overfitting results from complex model designs constructed using limited data, making them less universal²⁷. Explainable ML approaches have been developed based on deep learning and have a mesh-like structure that reduces overfitting to address these issues^{28,29}. Moreover, we aimed the development of a simple and accurate predictive model for COVID-19 severity for ease of use in clinical settings. To develop simple models, we combined reinforcement learning with explainable ML because the number of combinations of factors and operators is huge.

Materials and methods

Study Design and Settings

This retrospective cohort study used data from the Japan COVID-19 Task Force database collected between February 2020 and October 2022. The Japan COVID-19 Task Force collected clinical information on patients aged ≥ 18 years diagnosed with COVID-19 via polymerase chain reaction (PCR) or antigen testing at four institutions in Japan. Among the 3,424 patients identified, 123 participants, comprising those who lacked clinical data ($N=18$), were not positive by PCR or antigen test results ($N=5$), or were asymptomatic ($N=100$), were excluded. Finally, data from only 3,301 patients were analyzed. The patients were divided into discovery and validation cohorts. The discovery cohort comprised patients with estimated disease onset before October 1, 2020 ($N=1,023$), whereas the validation cohort comprised the remaining patients ($N=2,278$; Fig. 1). This study was approved by the Ethics Committee of Keio University School of Medicine (ID: 20200061) and adhered to the 1964 Declaration of Helsinki and its subsequent amendments. Written or oral informed consent was obtained from all the patients.

Clinical data

The following patient data were obtained from the electronic case record form: age, sex, body mass index (BMI), length of hospital stay (days), comorbidities, clinical signs and symptoms, laboratory findings, and post-hospitalization complications. All laboratory tests were performed within 48 h of the initial visit or admission based on the clinical care needs of the patients. A critical outcome was defined as the requirement for oxygen supplementation (via high-flow oxygen therapy), invasive or noninvasive mechanical ventilation, extracorporeal membrane oxygenation, or death^{4,30}.

Extraction of key features via ML

Predictive models were constructed for patients with and without critical outcomes to determine the features relevant to critical outcomes, and the importance of each feature was evaluated (Supplementary Table 1). Two baseline predictive models were constructed: a point-wise linear (PWL) model²⁷ (using PyTorch 1.5.1, Python 3.7.4), and a logistic regression model (using scikit-learn v0.24.2, Python 3.7.4). The output values of the PWL models were expressed as weighted sums of the input features, similar to logistic regression models. However, the weights of the PWL were calculated as nonlinear functions of the features using deep neural networks (in

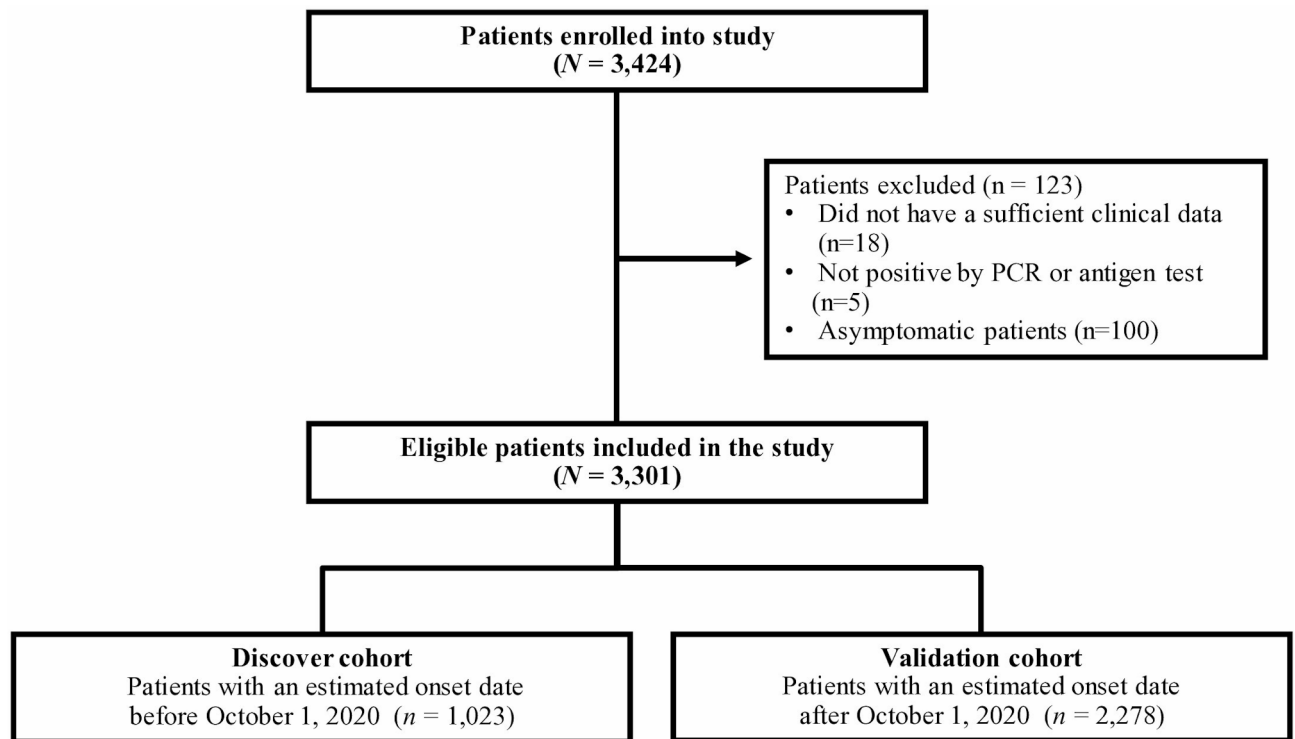


Fig. 1. Flowchart of the patient selection process. Flowchart of the patient identification and selection process. The 3,301 patients included were divided into the discovery and validation cohorts.

contrast to the logistic regression approach). This approach assigned different mathematical weights to each feature in the final algorithm.

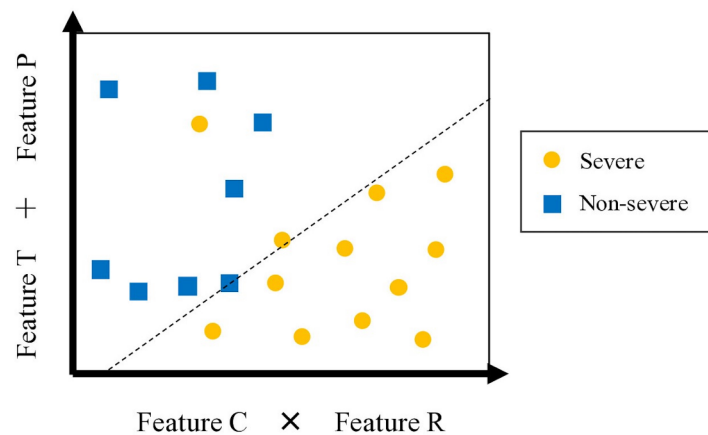
The feature variables were classified as binary (e.g., sex and pregnancy), multicategorical (e.g., blood type), or quantitative (e.g., Alb and BUN). All binary features were assigned a value of 1 or −1. One-hot encoding was used to convert multicategorical variables with K categories into K binary variables. Each variable was assigned a value of 1 if the sample belonged to the corresponding category; otherwise, a value of −1 was assigned. Each continuous quantitative variable was normalized by subtracting the mean value from the original value and dividing the result by the standard deviation (SD) subsequently. Missing data were assigned a value of zero because they did not change the output in the weighted-sum layers of the deep learning model. A total of 87 feature variables were generated.

The predictive performance of each model was calculated using the area under the receiver operating characteristic (ROC) curve (AUC) and evaluated using 10-fold double cross-validation (10-fold DCV) to avoid the dependencies on specific test sets. The probability thresholds for calculating the specificity and sensitivity of each model were determined by locating the point nearest to the point (0–1.0) on the ROC curve. The predictive performance of each model in the validation cohort was calculated using the model fitted by the discovery cohort dataset and the fold hyperparameter values with the best AUC values of the test set during the 10-fold DCV. The best hyperparameter values were represented in Supplementary Tables 2 and 3.

Extracting important features via ML and constructing predictive factors

A simple predictive model was constructed using the important features extracted from the deep learning model. The model specifications are shown in Fig. 2A. The logistic regression model is based on two predictive factors. Predictive factors were defined as those that could distinguish critical outcomes using significant features and basic mathematical operations, such as addition and square root functions. Table 1 shows the performance of the PWL and logistic regression models. The PWL model outperformed the logistic regression model in terms of the estimated value for all patient cohorts; however, their performances did not differ significantly, considering their error margins. The importance scores for each feature in both models are provided in Supplementary Tables 4, 5. Features with a relative score of >0.01 in the PWL were selected, as the PWL model demonstrated superior estimated performance and ability to consider nonlinear interactions. A flowchart of the analysis is represented in Fig. 2B. Ultimately, 41 features were extracted. Given the 2.66×10^{12} possible patterns for this simple predictive model (consisting of two predictors using 41 important features and 11 basic mathematical operations), reinforcement learning was used to generate a simple predictive model with high predictive performance. The classification boundaries of this model were set based on the threshold level determined by locating the point in the ROC curve nearest to the point (0–1.0). Further details of this process are provided in Supplementary Material (Supplementary Figs. 1, 2). The model with a high AUC in the discovery cohort was adapted for the validation cohort.

A



B

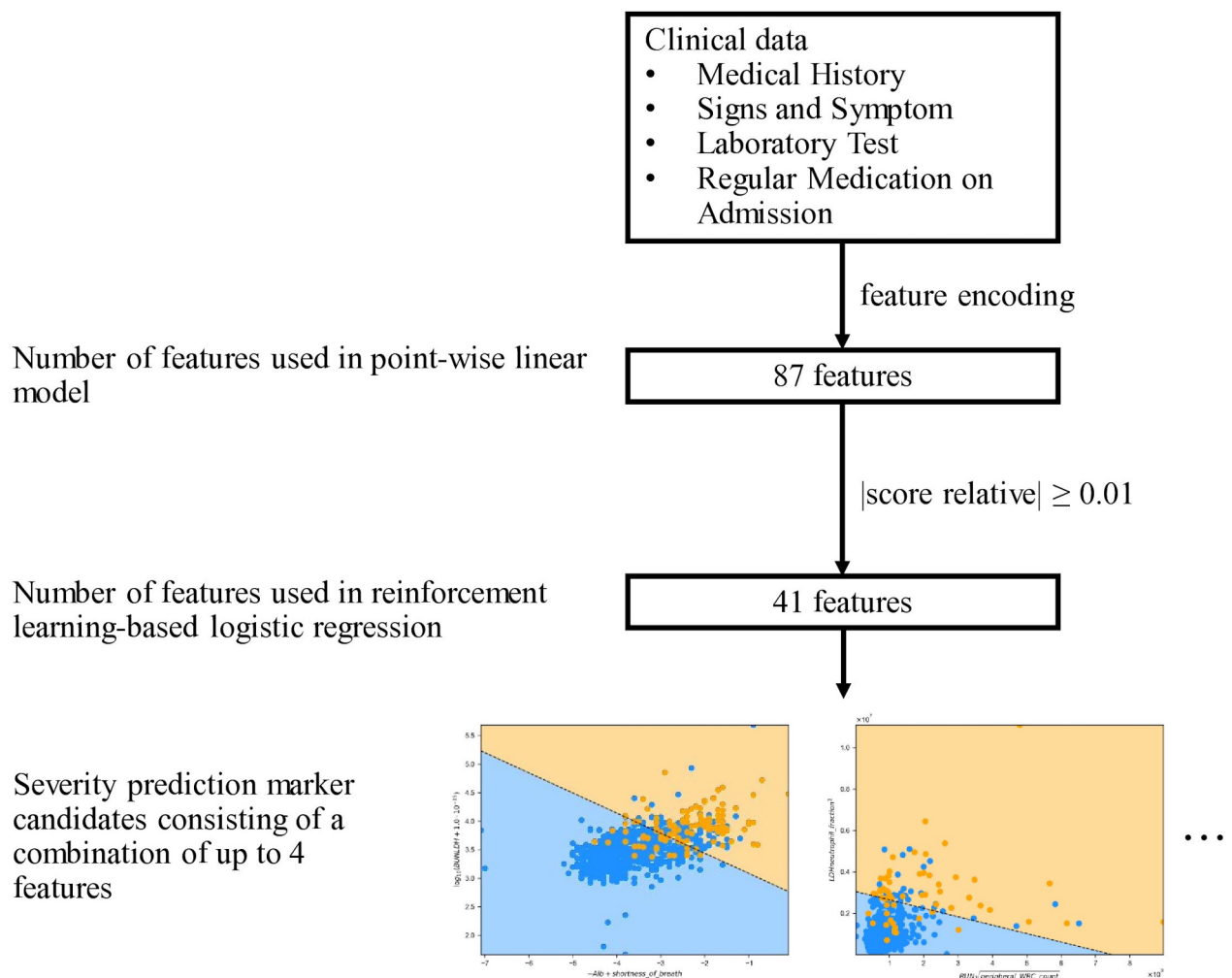


Fig. 2. Specifications of the simple predictive model for coronavirus disease 2019 severity and the flowchart of the analysis conducted in this study. (A) A simple predictive model for predicting critical COVID-19 was defined as a logistic regression model consisting of two predictive factors. A predictive factor was defined as a factor that could be used to distinguish critical illness based on extracted important features and basic mathematical operations. (B) The models were constructed in three steps. First, the clinical data were encoded to generate 87 features, according to a previously described approach. Second, predictive models were generated for COVID-19 severity using our PWL model—which extracted 41 features with a relative score of > 0.01 —to extract important features. The model was constructed based on these 41 features and basic mathematical operations but ultimately incorporated a combination of up to four features. COVID-19: coronavirus disease 2019, PWL: point-wise linear.

Models	Discovery Cohort						Validation Cohort		
	Training			Test					
	AUC	Specificity	Sensitivity	AUC	Specificity	Sensitivity	AUC	Specificity	Sensitivity
PWL	0.917 ± 0.005	0.892 ± 0.026	0.858 ± 0.012	0.915 ± 0.036	0.882 ± 0.038	0.792 ± 0.107	0.865	0.882	0.651
LR	0.912 ± 0.004	0.857 ± 0.019	0.854 ± 0.029	0.904 ± 0.038	0.830 ± 0.062	0.785 ± 0.110	0.864	0.878	0.644

Table 1. Prediction performance of the two models. PWL: point-wise linear model, LR: logistic regression with elastic-net regularization.

	Total	Discovery cohort (N= 1023)			Validation cohort (N= 2278)		
		Non-critical outcome	Critical outcome	P-value	Non-critical outcome	Critical outcome	P-value
		(N= 884)	(N= 139)		(N= 1823)	(N= 455)	
Age	56.4 ± 17.5	52.0 ± 19.4	65.4 ± 11.6	< 0.0001	56.1 ± 16.9	63.5 ± 13.4	< 0.0001
Female, %	32.8	37.8	24.5	0.0023	32.9	25.3	0.0018
BMI	24.8 ± 4.8	24.2 ± 4.6	25.7 ± 6.1	0.0021	24.8 ± 4.8	25.8 ± 4.7	0.0028
Smoking history, %	46.8	47.5	54.9	0.1263	44.8	51.5	0.0157
Smoking status (Brinkman index)	698.0 ± 2851.9	534.7 ± 847.4	769.8 ± 973.7	0.0519	779.9 ± 3814.2	664.8 ± 630.4	0.6664
Hypertension, %	33.6	25.2	50.4	< 0.0001	32.0	51.6	< 0.0001
Diabetes, %	21.1	16.7	37.2	< 0.0001	18.5	35.2	< 0.0001
Cardiovascular disease, %	10.1	7.2	20.1	< 0.0001	10.0	13.1	0.055
Cancer, %	6.4	4.3	9.4	0.0104	7.1	6.7	0.7914
Chronic obstructive pulmonary disease, %	4.1	3.3	9.6	0.0008	3.0	8.2	< 0.0001
Autoimmune disease, %	3.9	3.5	3.6	0.9208	4.4	2.9	0.1469
Asthma, %	7.3	6.5	5.2	0.583	7.9	7.4	0.7067
Hyperuricemia, %	10.1	7.4	17.3	0.0002	10.2	13.0	0.0808
Chronic liver disease, %	4.5	5.1	5.2	0.9902	3.9	5.1	0.2488
Chronic kidney disease, %	6.9	4.3	15.6	< 0.0001	5.2	15.5	< 0.0001

Table 2. Baseline characteristics of the patients.

Statistical analysis

For the baseline variables, categorical variables are expressed as frequencies and proportions, whereas continuous variables are expressed as means with SDs. Data were compared using the chi-square test for categorical variables and the Student’s t-test for continuous variables.

Results

Baseline characteristics of the patients

The baseline characteristics of the patients are summarized in Table 2. Among the 1,023 patients in the discovery cohort, the mean ages were 52.0 and 65.4 years, and 37.8% and 24.5% were women in the non-critical and critical outcome groups, respectively. The age and number of male patients were higher in the critical outcome group. A comparison of clinical characteristics revealed that BMI and incidence rates of hypertension, diabetes, cardiovascular disease, cancer, chronic obstructive pulmonary disease (COPD), hyperuricemia, and chronic kidney disease (CKD) were significantly higher in the critical outcome group. Among the 2,278 patients in the validation cohort, the mean ages were 56.1 and 63.5 years, and 32.9% and 25.3% were women in the non-critical and critical outcome groups, respectively. As observed in the discovery cohort, the age and number of men were higher in the critical outcome groups (BMI, hypertension, diabetes, COPD, and CKD). Unlike the discovery cohort, the validation cohort showed significant differences in terms of smoking history but no differences in terms of the prevalence of cardiovascular disease, cancer, and hyperuricemia.

Construction of a simple predictive model for COVID-19 severity using the extracted features

Using 41 features, the predictive factors were defined using reinforcement learning. The final model was defined as a two-dimensional logistic regression model that combined two prediction factors. The relationship between the model performance and the features with AUC values of ≥ 0.8 in the discovery cohort, assessed by reinforcement learning, is illustrated in Fig. 3. One of the models had an AUC of ≥ 0.905 and incorporated Alb, LDH, age, and neutrophil count. The other model, with an AUC of ≥ 0.800, incorporated 41 features. This finding suggests that the Alb and LDH levels, age, and neutrophil count are the most critical factors for achieving a high AUC.

The three simple predictive models with the highest AUC values for the discovery cohort are shown in Fig. 4A, and those for the validation cohort are shown in Fig. 4B. The top-performing model achieved AUC values of 0.906 in the discovery cohort (sensitivity, 0.842; specificity, 0.811; positive likelihood ratio, 4.456;

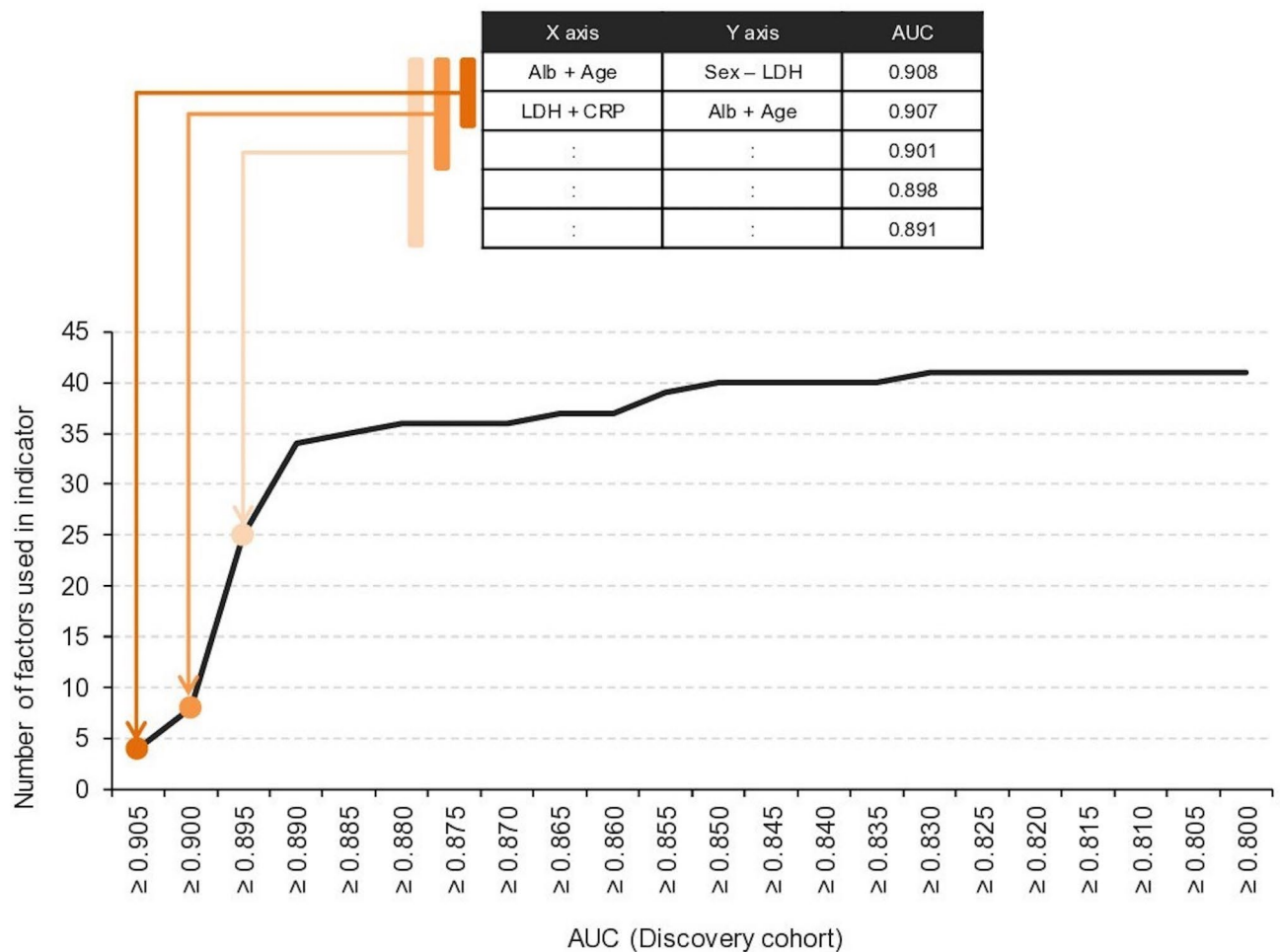


Fig. 3. Changes in the number of features used in the simple predictive model. This figure shows the relationship between the performance of the simple predictive models and the features used. The x-axis represents the “area under the receiver operating characteristic curve (AUC) ≥ 0.80 ,” which summarizes the number of unique features in the discovery cohort with an AUC value of ≥ 0.80 that were incorporated into the model. From left to right, the x-axis shows a wider range of AUC values. Thus, the number of simplified models increases. AUC: area under the receiver operating characteristic curve.

negative likelihood ratio, 0.195) and 0.861 in the validation cohort (sensitivity, 0.804; specificity, 0.675; positive likelihood ratio, 2.477; negative likelihood ratio, 0.290). This model utilizes only two predictors according to the following formula:

$$\text{Prediction factor 1} = -\sqrt{\text{LDH}} \times (\text{Neutrophil fraction})$$

$$\text{Prediction factor 2} = \frac{\text{Albumin}}{(\log_{10}(\text{Age} + 10^{-15}))^2}$$

Patients were predicted to experience critical COVID-19 when the following requirements were met:

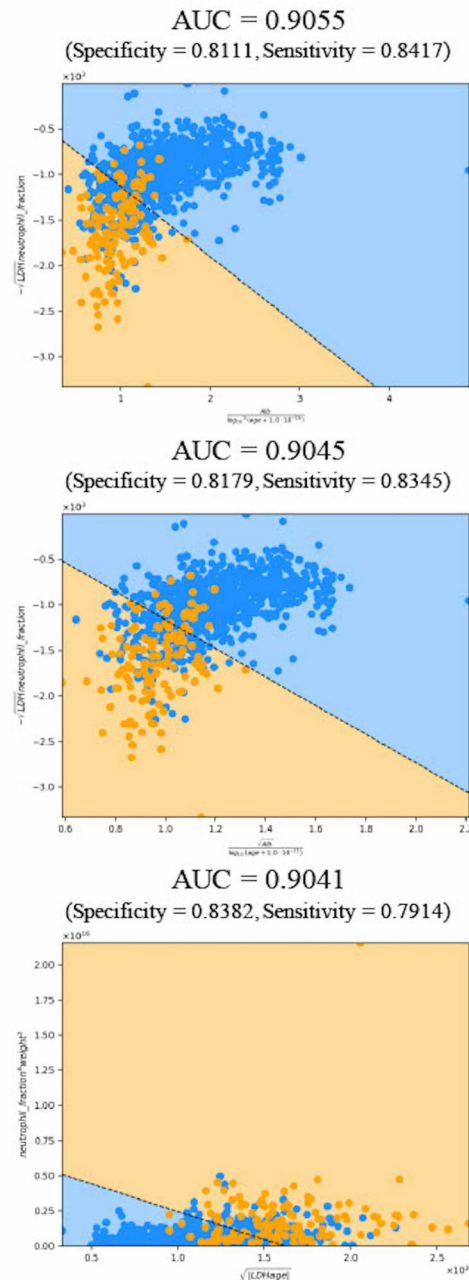
$$\text{Prediction factor 2} \leq -770.6 \times (\text{Prediction factor 1}) - 362.9$$

The second highest-performing model had AUC values of 0.905 in the discovery cohort (sensitivity, 0.835; specificity, 0.818; positive likelihood ratio, 4.583; negative likelihood ratio, 0.202) and 0.862 in the validation cohort (sensitivity, 0.785; specificity, 0.694; positive likelihood ratio, 2.563; negative likelihood ratio, 0.310). Moreover, it incorporates only two predictors according to the following formula:

$$\text{Prediction factor 1} = -\sqrt{\text{LDH}} \times (\text{Neutrophil fraction})$$

$$\text{Prediction factor 2} = \frac{\sqrt{\text{Albumin}}}{\log_{10}(\text{Age} + 10^{-15})}$$

A. Discovery Cohort



B. Validation Cohort

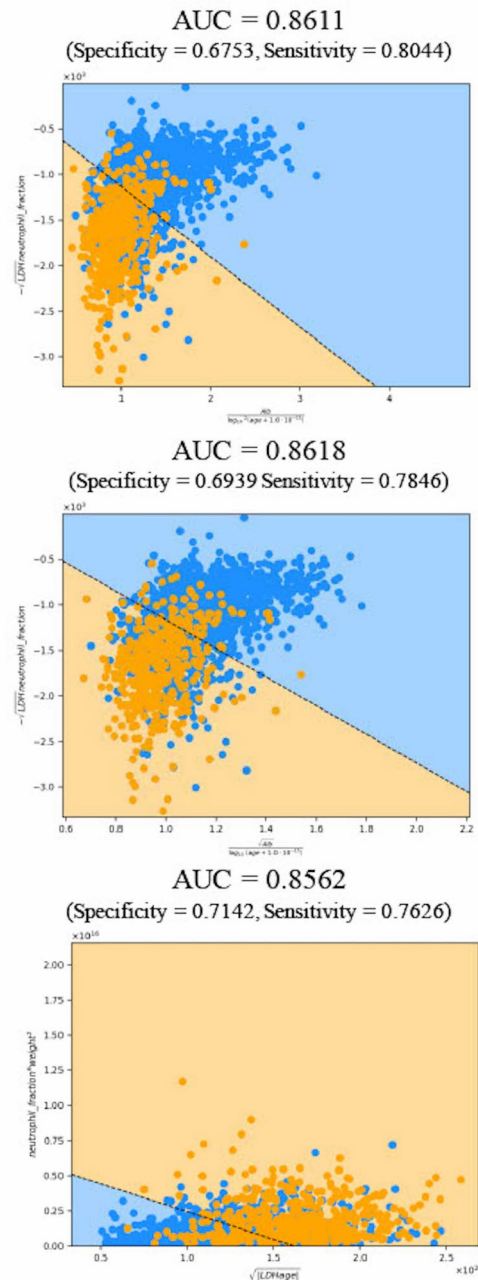


Fig. 4. Three simple predictive models with the highest area under the receiver operating characteristic curve values. The three simple predictive models with the highest AUC values in the discovery cohort. The orange and blue points correspond to patients with critical and noncritical outcomes, respectively. The orange and blue areas show the regions classified as critical and noncritical outcomes, respectively, for each model. The dashed line represents the classification boundaries obtained by setting the threshold level determined by locating the point in the ROC curve nearest to the point (0–1.0). (A) AUC values and plots for the discovery cohort. (B) AUC and plots for the validation cohort. AUC: area under the receiver operating characteristic curve.

Patients were predicted to experience critical COVID-19 when the following requirements were met:

$$\text{Prediction factor 1} \leq -1569.6 \times (\text{Prediction factor 2}) + 402.8$$

The third-highest-performing model had an AUC value of 0.904 in the discovery cohort (sensitivity, 0.791; specificity, 0.838; positive likelihood ratio, 4.891; negative likelihood ratio, 0.249) and 0.856 in the validation

cohort (sensitivity, 0.763; specificity, 0.714; positive likelihood ratio, 2.668; negative likelihood ratio, 0.332), which incorporated only two predictors according to the following formula:

$$\text{Prediction factor 1} = (\text{Neutrophil fraction})^6 (\text{weight})^2$$

$$\text{Prediction factor 2} = \sqrt{|\text{LDH} \times \text{Age}|}$$

Patients were predicted to experience critical COVID-19 when the following requirements were met:

$$\text{Prediction factor 1} \geq -3.95 \times 10^{13} (\text{Prediction factor 2}) + 6.40 \times 10^{15}$$

Discussion

This study developed an explanatory ML-based predictive model that utilized baseline characteristics and blood test findings of patients hospitalized with COVID-19. Critical cases were predicted using three models, with AUC values of 0.9055, 0.9045, and 0.9041. Age, LDH and Alb levels, and neutrophil count were the most significant factors. The reproducibility was confirmed using a validation cohort at different time points. This is the first study to establish an explanatory ML-based predictive model for COVID-19 severity that avoids the risk of overfitting and has high reproducibility.

Two predictive models were created to divide the patient cohort into patients with and without critical COVID-19 outcomes. The PWL model predicts the output using all the features and assigns an importance score to each feature. Conventional explainable deep learning models can generate importance scores; however, the importance score does not necessarily reveal the mechanism underlying the development of critical illness (i.e., how each feature contributes to the onset of critical illness). The main advantage of our approach is the construction of ML models of the second type, specifically, two-variable linear models, which incorporate at most two patient features related to elementary mathematical operations. These models elucidate the collaborative interactions between features in a highly interpretable manner. However, the total number of potential combinations for these models is vast: $(7 \times 4 \times 7 \times N \times 7 \times N)^2 - 2,000,000 \times N^4$, for “N” candidate features. To manage this complexity, the number of candidate features was initially limited to 50. Features were selected based on the importance scores assigned to the PWL model. A reinforcement learning method is then used to explore different model combinations to achieve high accuracy.

The PWL model was designed to avoid overfitting the results to the training data because it uses a unified architecture wherein each network layer and its associated neurons are interconnected in a mesh-like form³¹. Moreover, our conclusive models were two-variable linear models, each of which incorporated a limited number of features and mathematical operations. Such models have a lower risk of overfitting because of the smaller parameter count compared with the number of patients. We developed simple prediction models using the important features extracted from PWL model by reinforcement learning. The best prediction performance of the simple prediction model was the AUC values of 0.906 and 0.861 in discovery and validation cohort, respectively (Fig. 4B). Those results surpass the performance of a severity prediction model that incorporated commodities and laboratory data constructed using machine learning (AUC values of 0.786)³². In addition, the performance of the severe prediction model that incorporated six factors extracted by the statistical approach were the AUC values of 0.86 and 0.83 in derivation and validation population, respectively³³. In contrast, the simple prediction model developed in this study achieves high predictive performance using four features, suggesting that it can serve as a method to achieve a high level of accuracy while minimizing the number of required predictors. Notably, the model in this study achieves comparable or superior predictive accuracy despite using fewer factors. These findings suggest that, in severity prediction, modeling the interactions among factors appropriately may contribute more to improving predictive accuracy than merely increasing the number of factors.

A major strength of this study is its high reproducibility in a validation cohort of patients from different waves of the COVID-19 pandemic in Japan. Variations were observed in the clinical characteristics of hospitalized patients with COVID-19 during different periods, as documented in our previous study³⁴. These variations occur because of a variety of factors, such as the viral strain, medical environment, and laboratory system³⁵. The high predictive power of our model in both cohorts suggests its significant clinical utility. However, the AUC did not consistently surpass that of previous similar AI-based models^{32,36}. This may be attributable to the comprehensive analysis of a large patient group over a long period. Another strength is the simplicity of our approach, which focuses on common indices, such as patient demographics, LDH, and BUN. Many AI-based predictive models incorporate radiographic and computed tomography indexes³⁷; however, our models are straightforward and cost-effective. Our previous study highlighted differences in genetic characteristics, such as polymorphisms and blood groups³⁸, comorbidities, such as obesity⁷, and clinical characteristics, such as disease severity, between patients with COVID-19 in Japan and those in Western countries³⁹. Consequently, predictive models for COVID-19 severity are particularly important for Japanese patients. Although several predictive models have been established in Japan, most are primarily based on single-center data^{33,40}, with only one ML-based study incorporating multicenter data⁴¹. In practice, models developed by patient data obtained at a single institution often exhibited diminished performance in external validation³². However, the model developed in this study was constructed using multicenter data from Japan. Therefore, it is less likely to experience performance degradation when applied to new Japanese patient populations. Therefore, a significant strength of our study is the development of a predictive model for COVID-19 severity in Japanese patients with COVID-19 using comprehensive multicenter data.

Age, neutrophil count, LDH levels, and Alb levels were selected as factors in our predictive model. Age is a significant prognostic factor for COVID-19^{42,43}, and its fatality rate increases with age⁴⁴. A meta-analysis examining the sole effect of age reported that age increased the risk of hospitalization, in-hospital mortality, and

COVID-19-related mortality⁴⁵. This association is attributed to the tendency for the prevalence of comorbidities that exacerbate disease severity to increase with age; however, other factors have also been reported. Older adults often experience impaired clearance of dead cells, persistent inflammation, and heightened secretion of inflammatory molecules, which contribute to the exacerbation of the cytokine storm (a typical characteristic of COVID-19)⁴⁶. Indeed, the association between NLRP3 inflammasome and COVID-19 severity have been reported. NLRP3 activity is associated with cytokine storm that results in severity⁴⁷, and NLRP3 is over-activated in aged patients⁴⁸. Age is known as a risk factor for various disease, not only COVID-19, and may be related to metabolism and resistance to treatment as well as comorbidities and abnormal immune response. With aging, the immune system becomes weakened, and immune regulation fails, which may lead to severe respiratory failure in COVID-19. Other factors such as age-related malnutrition, sarcopenia, changes in the respiratory system, and hormonal changes have also been suspected⁴⁹, and these factors are thought to interact and cause the severity of disease.

COVID-19 causes cytokine storms, and increased neutrophil counts can be attributed to hyperinflammatory conditions. This increase may have been exacerbated by concomitant bacterial infections. Severe COVID-19 patients had higher neutrophil counts at admission than mild or moderate patients⁵⁰, and elevated neutrophil counts was reported to be an indicator of severity⁵¹. Severe cases of COVID-19 are associated with high expression of neutrophil-associated cytokines, and neutrophilia has been identified as a predictor of poor prognosis⁵². Several mechanisms contribute to this, including higher levels of low-density neutrophils (LDNs) and the formation of neutrophil extracellular traps (NETs)^{53,54}. LDN levels are particularly elevated in patients with severe COVID-19⁵⁵ and are associated with NET formation, which contributes to lung injury and vascular occlusion⁵⁶. Elevated neutrophil activation promotes NET formation, which activates the coagulation cascade and contributes to microangiopathy and thrombosis^{57,58}. In proteomic analysis, this was associated with an increase in proteins involved in metabolism, immunosuppression, and pattern recognition in severe patients⁵⁹. These mechanisms of neutrophils were suggested to be related with the severity.

LDH is an intracellular enzyme found in various organs that oxidizes pyruvate to lactate. LDH is commonly used as a marker of tissue damage in various diseases and is associated with lung damage and pneumonia. LDH is a terminal enzyme involved in anaerobic glycolysis, and elevated LDH levels indicate oxygen deficiency. LDH is associated with severity of disease in COVID-19⁶⁰ and is indicated as a factor associated with the extent of COVID-19 pneumonia by computed tomography⁶¹. In the meta-analysis, LDH was higher in patients who died compared to those who survived, and in patients with severe disease compared to those with non-severe disease⁶². LDH levels increase in response to tissue damage, promote the synthesis of lactate, and elevate the levels of cells, such as macrophages and dendritic cells, while suppressing natural killer cells and cytotoxic T lymphocytes⁶³, potentially exacerbating COVID-19 severity.

Hypoalbuminemia is a well-known predictor of disease severity. A previous meta-analysis reported that it was associated with a poor prognosis⁶⁴. The mechanism underlying the association between COVID-19 and hypoalbuminemia remains unknown; however, it is thought to increase vascular permeability and shorten the half-life of Alb through inflammation. Another study reported an inverse relationship between serum Alb levels and venous thromboembolism (VTE) in critically ill patients⁶⁵. Lower serum Alb levels reduced the incidence of VTE in patients with COVID-19⁶⁶. This underscores the complex interplay between hypercoagulability-induced hypoalbuminemia and neutrophil hypercoagulability, which potentially exacerbates disease severity.

There were several limitations in this study. First, an internal validation was performed for this study. However, external validation, particularly in patient populations from other regions, has not been performed. There are known ethnical differences in hospitalization and mortality rates of COVID-19⁶⁷. Therefore, large-scale studies should be conducted in other countries to confirm the reproducibility of these results. Second, the study duration was relatively short. It was reported that the severity of COVID-19 varied depending on strain. Omicron variant was reported to cause milder COVID-19 than earlier strains^{68,69}, and to be more contagious⁷⁰. Therefore, research of COVID-19 should take into account the prevalent strain. Because the most recent Omicron strain has not been sufficiently investigated, validation studies at different time points are warranted. Furthermore, various treatments for COVID-19 have emerged over time⁷¹, necessitating further studies to evaluate these new approaches. Third, data on vaccinations were not accessible. Vaccination was reported to reduce the severity of COVID-19 in systematic review⁷² and similar effects was also shown in Japan⁷³, where vaccine administration is recommended by the national policies. Therefore, the applicability of our predictive model in vaccinated patients should be examined.

In this multicenter study, simple and well-structured predictive models for COVID-19 severity in Japanese patients were established using explanatory ML. These results may aid in patient management and the selection of effective therapeutic interventions.

Data availability

The datasets used and/or analyzed in the current study are available from the corresponding author upon reasonable request.

Received: 7 September 2024; Accepted: 6 January 2025

Published online: 19 March 2025

References

1. Zhou, P. et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **579**, 270–273 (2020).
2. Hu, B., Guo, H., Zhou, P. & Shi, Z. L. Characteristics of SARS-CoV-2 and COVID-19. *Nat. Rev. Microbiol.* **19**, 141–154 (2021).
3. Yang, Z. R. et al. Efficacy of SARS-CoV-2 vaccines and the dose-response relationship with three major antibodies: a systematic review and meta-analysis of randomised controlled trials. *Lancet Microbe.* **4**, e236–e246 (2023).

4. Tanaka, H. et al. Characteristics and clinical effectiveness of COVID-19 vaccination in hospitalized patients in Omicron-dominated epidemic wave – a nationwide study in Japan. *Int. J. Infect. Dis.* **132**, 84–88 (2023).
5. Yuan, Y., Jiao, B., Qu, L., Yang, D. & Liu, R. The development of COVID-19 treatment. *Front. Immunol.* **14**, 1125246 (2023).
6. Tanaka, H. et al. Propensity-score matched analysis of the effectiveness of baricitinib in patients with coronavirus disease 2019 (COVID-19) using nationwide real-world data: an observational matched cohort study from the Japan COVID-19 task force. *Open. Forum Infect. Dis.* **10**, 311 (2023).
7. Lee, H. et al. Effects of mild obesity on outcomes in Japanese patients with COVID-19: a nationwide consortium to investigate COVID-19 host genetics. *Nutr. Diabetes.* **12**, 38 (2022).
8. Fukushima, T. et al. Clinical significance of prediabetes, undiagnosed diabetes and diagnosed diabetes on critical outcomes in COVID-19: integrative analysis from the Japan COVID-19 task force. *Diabetes Obes. Metab.* **25**, 144–155 (2023).
9. Fukushima, T. et al. U-shaped association between abnormal serum uric acid levels and COVID-19 severity: reports from the Japan COVID-19 Task Force. *Int. J. Infect. Dis.* **122**, 747–754 (2022).
10. Azekawa, S. et al. Serum KL-6 levels predict clinical outcomes and are associated with MUC1 polymorphism in Japanese patients with COVID-19. *BMJ Open. Respir Res.* **10**, e001625 (2023).
11. Zhu, B. et al. Correlation between white blood cell count at admission and mortality in COVID-19 patients: a retrospective study. *BMC Infect. Dis.* **21**, 574 (2021).
12. Chi, L., Wang, S., Wang, X., Yang, C. & Luo, J. Predictive value of C-reactive protein for disease severity and survival in COVID-19 patients: a systematic review and meta-analysis. *Clin. Exp. Med.* **23**, 2001–2008 (2023).
13. Sharma, J., Rajput, R., Bhatia, M., Arora, P. & Sood, V. Clinical predictors of COVID-19 severity and mortality: a perspective. *Front. Cell. Infect. Microbiol.* **11**, 674277 (2021).
14. Henry, B. M. et al. Lactate dehydrogenase levels predict coronavirus disease 2019 (COVID-19) severity and mortality: a pooled analysis. *Am. J. Emerg. Med.* **38**, 1722–1726 (2020).
15. Chen, X. Y., Huang, M. Y., Xiao, Z. W., Yang, S. & Chen, X. Q. Lactate dehydrogenase elevations is associated with severity of COVID-19: a meta-analysis. *Crit. Care.* **24**, 459 (2020).
16. Alizadeh, N. et al. Lactate dehydrogenase to albumin ratio as a predictive factor of COVID-19 patients' outcome; a cross-sectional study. *Arch. Acad. Emerg. Med.* **10**, e63 (2022).
17. Shokr, H. et al. Lactate dehydrogenase/albumin to-urea ratio: a novel prognostic maker for fatal clinical complications in patients with COVID-19 infection. *J. Clin. Med.* **12** (1), 19 (2022).
18. Li, X. et al. Predictive values of neutrophil-to-lymphocyte ratio on disease severity and mortality in COVID-19 patients: a systematic review and meta-analysis. *Crit. Care.* **24**, 647 (2020).
19. Zarębska-Michaluk, D. et al. Impact of kidney failure on the severity of COVID-19. *J. Clin. Med.* **10**, 2042 (2021).
20. Zhang, J. et al. Liver fibrosis scores and clinical outcomes in patients with COVID-19. *Front. Med. (Lausanne)* **9**, 829423 (2022).
21. Wong, J. J. M. et al. Development and validation of a clinical predictive model for severe and critical pediatric COVID-19 infection. *PLoS ONE* **17**, e0275761 (2022).
22. Lee, J. Y. et al. A risk scoring system to predict progression to severe pneumonia in patients with COVID-19. *Sci. Rep.* **12**, 5390 (2022).
23. Trujillo-Rodriguez, M. et al. Clinical, laboratory data and inflammatory biomarkers at baseline as early discharge predictors in hospitalized SARS-CoV-2 infected patients. *PLoS ONE* **17**, e0269875 (2022).
24. Shashikumar, S. P. et al. Development and prospective validation of a deep learning algorithm for predicting need for mechanical ventilation. *Chest* **159**, 2264–2273 (2021).
25. Liu, Q. et al. Machine learning models for predicting critical illness risk in hospitalized patients with COVID-19 pneumonia. *J. Thorac. Dis.* **13**, 1215–1229 (2021).
26. Charilaou, P. & Battat, R. Machine learning models and over-fitting considerations. *World J. Gastroenterol.* **28**, 605–607 (2022).
27. Parmar, C., Barry, J. D., Hosny, A., Quackenbush, J. & Aerts, H. J. W. L. Data analysis strategies in medical imaging. *Clin. Cancer Res.* **24**, 3492–3499 (2018).
28. Hirakawa, Y. et al. Potential progression biomarkers of diabetic kidney disease determined using comprehensive machine learning analysis of non-targeted metabolomics. *Sci. Rep.* **12**, 16287 (2022).
29. Kumagai, S. et al. The PD-1 expression balance between effector and regulatory T cells predicts the clinical efficacy of PD-1 blockade therapies. *Nat. Immunol.* **21**, 1346–1358 (2020).
30. Otake, S. et al. Clinical clustering with prognostic implications in Japanese COVID-19 patients: report from Japan COVID-19 Task Force, a nation-wide consortium to investigate COVID-19 host genetics. *BMC Infect. Dis.* **22**, 735 (2022).
31. Golas, S. B. et al. A machine learning model to predict the risk of 30-day readmissions in patients with heart failure: a retrospective analysis of electronic medical records data. *BMC Med. Inf. Decis. Mak.* **18**, 44 (2018).
32. Jimenez-Solem, E. et al. Developing and validating COVID-19 adverse outcome risk prediction models from a bi-national European cohort of 5594 patients. *Sci. Rep.* **11**, 3246 (2021).
33. Muto, Y. et al. Predictive model for the development of critical coronavirus disease 2019 and its risk factors among patients in Japan. *Respir Investig.* **59**, 804–809 (2021).
34. Lee, H. et al. Characteristics of hospitalized patients with COVID-19 during the first to fifth waves of infection: a report from the Japan COVID-19 Task Force. *BMC Infect. Dis.* **22**, 935 (2022).
35. Dave, N. et al. Nosocomial SARS-CoV-2 infections and mortality during unique COVID-19 epidemic waves. *JAMA Netw. Open.* **6**, e2341936 (2023).
36. Yadaw, A. S. et al. Clinical features of COVID-19 mortality: development and validation of a clinical prediction model. *Lancet Digit. Health.* **2**, e516–e525 (2020).
37. Adamidi, E. S., Mitsis, K. & Nikita, K. S. Artificial intelligence in clinical care amidst COVID-19 pandemic: a systematic review. *Comput. Struct. Biotechnol. J.* **19**, 2833–2850 (2021).
38. Kusumoto, T. et al. Association between ABO blood group/genotype and COVID-19 in a Japanese population. *Ann. Hematol.* **102**, 3239–3249 (2023).
39. Tanaka, H. et al. Clinical characteristics of patients with coronavirus disease (COVID-19): preliminary baseline report of Japan COVID-19 task force, a nationwide consortium to investigate Host Genetics of COVID-19. *Int. J. Infect. Dis.* **113**, 74–81 (2021).
40. Kitajima, H. et al. Scoring system for identifying Japanese patients with COVID-19 at risk of requiring oxygen supply: a retrospective single-center study. *J. Infect. Chemother.* **27**, 1217–1222 (2021).
41. Nojiri, S., Irie, Y., Kanamori, R., Naito, T. & Nishizaki, Y. Mortality prediction of COVID-19 in hospitalized patients using the 2020 diagnosis procedure combination administrative database of Japan. *Intern. Med.* **62**, 201–213 (2023).
42. Verity, R. et al. Estimates of the severity of coronavirus disease 2019: a model-based analysis. *Lancet Infect. Dis.* **20**, 669–677 (2020).
43. Statsenko, Y. et al. Impact of age and sex on COVID-19 severity assessed from radiologic and clinical findings. *Front. Cell. Infect. Microbiol.* **11**, 777070 (2021).
44. Levin, A. T. et al. Assessing the age specificity of infection fatality rates for COVID-19: systematic review, meta-analysis, and public policy implications. *Eur. J. Epidemiol.* **35**, 1123–1138 (2020).
45. Romero Starke, K. et al. The isolated effect of age on the risk of COVID-19 severe outcomes: a systematic review with meta-analysis. *BMJ Glob Health.* **6**, e006434 (2021).
46. Akbar, A. N. & Gilroy, D. W. Aging immunity may exacerbate COVID-19. *Science* **369**, 256–257 (2020).

47. Pan, P. et al. SARS-CoV-2 N protein promotes NLRP3 inflammasome activation to induce hyperinflammation. *Nat. Commun.* **12**, 4664 (2021).
48. Lara, P. C. et al. Age-induced NLRP3 inflammasome over-activation increases lethality of SARS-CoV-2 Pneumonia in Elderly patients. *Aging Dis.* **11**, 756–762 (2020).
49. Gallo, A. et al. How can Biology of Aging explain the severity of COVID-19 in older adults. *Clin. Geriatr. Med.* **38**, 461–472 (2022).
50. Chen, R. et al. Longitudinal hematologic and immunologic variations associated with the progression of COVID-19 patients in China. *J. Allergy Clin. Immunol.* **146**, 89–100 (2020).
51. Chiang, C. C. et al. Targeting neutrophils to treat Acute Respiratory Distress Syndrome in Coronavirus Disease. *Front. Pharmacol.* **11**, 572009 (2020).
52. Vafadar Moradi, E. et al. Increased age, neutrophil-to-lymphocyte ratio (NLR) and white blood cells count are associated with higher COVID-19 mortality. *Am. J. Emerg. Med.* **40**, 11–14 (2021).
53. Morrissey, S. M. et al. A specific low-density neutrophil population correlates with hypercoagulation and disease severity in hospitalized COVID-19 patients. *JCI Insight.* **6**, e148435 (2021).
54. Pastorek, M., Dúbrava, M. & Celec, P. On the origin of neutrophil extracellular traps in COVID-19. *Front. Immunol.* **13**, 821007 (2022).
55. Manunta, M. D. I. et al. Impact of SARS-CoV-2 infection on the recovery of peripheral blood mononuclear cells by density gradient. *Sci. Rep.* **11**, 4904 (2021).
56. Obermayer, A. et al. Neutrophil extracellular traps in fatal COVID-19-associated lung injury. *Dis. Markers* 5566826 (2021). (2021).
57. Carsetti, R. et al. Different innate and adaptive immune responses to SARS-CoV-2 infection of asymptomatic, mild, and severe cases. *Front. Immunol.* **11**, 610300 (2020).
58. Bordallo, B. et al. COVID-19: what have we learned with the immunopathogenesis? *Adv. Rheumatol.* **60**, 50 (2020).
59. Long, M. B. et al. Extensive acute and sustained changes to neutrophil proteomes post-SARS-CoV-2 infection. *Eur. Respir J.* **63**, 2300787 (2024).
60. Li, X. et al. Risk factors for severity and mortality in adult COVID-19 inpatients in Wuhan. *J. Allergy Clin. Immunol.* **146**, 110–118 (2020).
61. Kojima, K., Yoon, H., Okishio, K. & Tsuyuguchi, K. Increased lactate dehydrogenase reflects the progression of COVID-19 pneumonia on chest computed tomography and predicts subsequent severe disease. *Sci. Rep.* **13**, 1012 (2023).
62. Fialek, B. et al. Diagnostic value of lactate dehydrogenase in COVID-19: a systematic review and meta-analysis. *Cardiol. J.* **29**, 751–758 (2022).
63. Gupta, G. S. The lactate and the lactate dehydrogenase in inflammatory diseases and major risk factors in COVID-19 patients. *Inflammation* **45**, 2091–2123 (2022).
64. Soetedjo, N. N. M. et al. Prognostic properties of hypoalbuminemia in COVID-19 patients: a systematic review and diagnostic meta-analysis. *Clin. Nutr. ESPEN.* **45**, 120–126 (2021).
65. Chi, G. et al. Inverse relationship of serum albumin to the risk of venous thromboembolism among acutely ill hospitalized patients: analysis from the APEX trial. *Am. J. Hematol.* **94**, 21–28 (2019).
66. Kheir, M., Saleem, F., Wang, C., Mann, A. & Chua, J. Higher albumin levels on admission predict better prognosis in patients with confirmed COVID-19. *PLoS ONE.* **16**, e0248358 (2021).
67. Irizar, P. et al. Ethnic inequalities in COVID-19 infection, hospitalisation, intensive care admission, and death: a global systematic review and meta-analysis of over 200 million study participants. *EClinicalMedicine* **57**, 101877 (2023).
68. Esper, F. P. et al. Alpha to Omicron: Disease Severity and Clinical outcomes of Major SARS-CoV-2 variants. *J. Infect. Dis.* **227**, 344–352 (2023).
69. Arabi, M. et al. Severity of the Omicron SARS-CoV-2 variant compared with the previous lineages: a systematic review. *J. Cell. Mol. Med.* **27**, 1443–1464 (2023).
70. Araf, Y. et al. Omicron variant of SARS-CoV-2: Genomics, transmissibility, and responses to current COVID-19 vaccines. *J. Med. Virol.* **94**, 1825–1832 (2022).
71. Li, G., Hilgenfeld, R., Whitley, R. & Clercq, E. D. Therapeutic strategies for COVID-19: progress and lessons learned. *Nat. Rev. Drug Discov.* **22**, 449–475 (2023).
72. Graña, C. et al. Efficacy and safety of COVID-19 vaccines. *Cochrane Database Syst. Rev.* **12**, CD015477 (2022).
73. Arashiro, T. et al. COVID-19 vaccine effectiveness against severe COVID-19 requiring oxygen therapy, invasive mechanical ventilation, and death in Japan: a multicenter case-control study (MOTIVATE study). *Vaccine* **42**, 677–688 (2024).

Acknowledgements

We thank all the patients involved in this study. We appreciate the support of all members of the Japan COVID-19 Task Force, who are regularly engaged in clinical and research work on COVID-19.

Author contributions

Conceptualization: T.O., S.C., H.N., and S.I. Data curation: H.T, H.L., T.F., S.A., S.O., K.N., and M.W. Formal analysis: T.O., S.C., H.N., S.N., R.I., and S.I. Funding acquisition: S.C., H.N., T.A., Y.O., M.I., T.K., K.F., and S.I. Methodology: T.O., S.C., H.N., and S.I. Supervision: K.M., H.K., N.H., T.U., S.U., T.I., K.A., F.S., T.Y., Y.N., Y.M., Y.S., R.E., K.M., Y.S., Y.O., R.K., M.I., N.H., Y.K., K.T., A.K., S.M., S.O., T.K., and K.F. Visualization: T.O., S.C., H.N., S.N., R.I., and S.I. Writing—original draft: T.O., S.C., H.N., and S.I. Writing—review and editing: T.A., K.M., H.K., N.H., T.U., S.U., T.I., K.A., F.S., T.Y., Y.N., Y.M., Y.S., R.E., K.M., Y.S., Y.O., R.K., M.I., N.H., Y.K., K.T., A.K., S.M., S.O., T.K., and K.F.

Funding

This work was supported by the Japan Agency for Medical Research and Development [grant numbers: JP20nk0101612, JP20fk0108415, JP21jk0210034, JP21km0405211, JP21km0405217, JP21wm0325031, JP20fk0108452, JP21fk0108563, JP21fk0108553, JP21fk0108573, JP22fk0108573, JP21fk0108431, JP22fk0108510, JP22wm0325031, and JP23tm0524008], Japan Science and Technology Agency CREST [grant number: JPM-JCR20H2], Japan Science and Technology Agency PRESTO [grant number: JPMJPR21R7], and the Ministry of Health, Labour and Welfare [grant number: 20CA2054].

Declarations

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/s41598-025-85733-5>.

Correspondence and requests for materials should be addressed to S.C., H.N. or S.I.

Reprints and permissions information is available at 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-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, 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 you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. 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-nc-nd/4.0/>.

© The Author(s) 2025