Heliyon 9 (2023) e18764

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Severity-onset prediction of COVID-19 via artificial-intelligence analysis of multivariate factors

Yu Fu [a,1], Lijiao Zeng [a,1], Pilai Huang [a,1], Mingfeng Liao [a,1], Jialu Li [d], Mingxia Zhang [a], Qinlang Shi [a], Zhaohua Xia [a], Xinzhong Ning [a], Jiu Mo [d], Ziyuan Zhou [a], Zigang Li [c], Jing Yuan [a], Lifei Wang [a], Qing He [a], Qikang Wu [e,**], Lei Liu [a,***], Yuhui Liao [b,*], Kun Qiao [a,****]

[a] Department of Infectious Diseases, Department of Thoracic Surgery, Department of Radiology, National Clinical Research Center for Infectious Disease, The Second Affiliated Hospital of Southern University of Science and Technology, Shenzhen Third People's Hospital, Shenzhen, China
[b] Molecular Diagnosis and Treatment Center for Infectious Diseases, Dermatology Hospital, Southern Medical University, Guangzhou, China
[c] Pingshan Translational Medicine Center, Shenzhen Bay Laboratory, and State Key Laboratory of Chemical Oncogenomics, School of Chemical Biology and Biotechnology, Peking University Shenzhen Graduate School, Shenzhen, China
[d] Department of Biostatistics, HuaJia Biomedical Intelligence, Shenzhen, China
[e] Department of Clinical Laboratory, The First People's Hospital of Foshan, Foshan, China

ARTICLE INFO

Keywords:
COVID-19
Severe disease onset prediction
Artificial-intelligence analysis

ABSTRACT

Progression to a severe condition remains a major risk factor for the COVID-19 mortality. Robust models that predict the onset of severe COVID-19 are urgently required to support sensitive decisions regarding patients and their treatments. In this study, we developed a multivariate survival model based on early-stage CT images and other physiological indicators and biomarkers using artificial-intelligence analysis to assess the risk of severe COVID-19 onset. We retrospectively enrolled 338 adult patients admitted to a hospital in China (severity rate, 31.9%; mortality rate, 0.9%). The physiological and pathological characteristics of the patients with severe and non-severe outcomes were compared. Age, body mass index, fever symptoms upon admission, coexisting hypertension, and diabetes were the risk factors for severe progression. Compared with the non-severe group, the severe group demonstrated abnormalities in biomarkers indicating organ function, inflammatory responses, blood oxygen, and coagulation function at an early stage. In addition, by integrating the intuitive CT images, the multivariable survival model showed significantly improved performance in predicting the onset of severe disease (mean time-dependent area under the curve = 0.880). Multivariate survival models based on early-stage CT images and other physiological indicators and biomarkers have shown high potential for predicting the onset of severe COVID-19.

* Corresponding authors.
** Corresponding author.
*** Corresponding authors.
**** Corresponding author.
E-mail addresses: wqkang@fsyyy.com (Q. Wu), liulei1961@hotmail.com (L. Liu), liaoyh8@mail.sysu.edu.cn (Y. Liao), szqiaokun@163.com (K. Qiao).
[1] These authors contributed equally.

https://doi.org/10.1016/j.heliyon.2023.e18764
Received 27 April 2023; Received in revised form 13 July 2023; Accepted 26 July 2023
Available online 27 July 2023

1. Introduction

To date, over 12 million people worldwide have been confirmed to have coronavirus disease (COVID-19) and more than 2.6 million people have died from the disease [1]. This high mortality rate has placed tremendous pressure on medical institutions worldwide. Progression into a severe condition is an important risk factor of death, especially for patients without obvious severe symptoms [2,3]. For such patients, owing to the lack of effective decision support tools, medical decisions are limited by the doctors' clinical experience, which makes them prone to miscalculations and may delay the best treatment opportunity [4]. To reduce the mortality rate of COVID-19, there is an urgent need for a pragmatic risk stratification tool that will allow the early identification of patients with the highest risk of severe COVID-19 onset to guide management and optimize resource allocation [5,6].

Previous studies explored risk factors associated with the disease severity [7,8]. For example, Zhou et al. [9] found that older age and a high D-dimer levels are associated with in-hospital death. Shi et al. [10] found that abnormalities in the manifestation of chest CT imaging correlate with disease states. In addition, Wang et al. [11] found that the percentage of pulmonary infection volume was larger in cases with hypertension and was a considerable risk factor for severe COVID-19 in patients with hypertension. Furthermore, Alberca et al. [12] clarified that obesity is a risk factor for the development of severe COVID-19 with the need for hospitalization and mechanical ventilation. Previous research findings have enhanced the understanding of the SARS-CoV-2 etiology and pathophysiology of the disease, but they only identified that a certain factor (older age, high D-dimer level, CT imaging, hypertension, or obesity) may lead to severe disease development in patients with COVID-19; however, there is no simultaneous analysis of multiple factors, which is limiting, and the early prediction of severe cases is poor. Nevertheless, comprehensive risk prediction of severe conversion may be the best way to reduce the COVID-19 mortality. Many clinical prediction models of COVID-19 have been developed; they vary in their setting, predicted outcome measures, and included clinical parameters, but most of them show moderate performance and no benefit to clinical decision-making. In addition, a joint analysis of the timing and progression of events and the integration of multiple types of input data into risk prediction have not been investigated thoroughly. In this study, we aimed to develop and validate a pragmatic prediction model for onset of severe disease based on early-stage CT images and other physiological indicators and biomarkers in patients without obvious severe symptoms at an early stage. We retrospectively collected data on a cohort that was characterized by a large proportion of imported cases and used these cohort data to develop a statistical model that can provide the real-time risk of severe disease onset for COVID-19 patients through an artificial-intelligence analysis of multivariate factors (Fig. 1).

With this strategy, 76 (31.9%) patients progressed to severe conditions, and 3 (0.9%) died. The mean time from hospital admission to onset of severe COVID-19 was 3.7 days. Age, body mass index (BMI), fever symptoms upon admission, coexisting hypertension, and diabetes were the risk factors for severe progression. Compared with the non-severe group, the severe group had abnormalities in biomarkers indicating organ function, inflammatory responses, blood oxygen and coagulation function at an early stage. The cohort was characterized by an increasing cumulative incidence of severe disease progression for up to 10 days after hospitalization. The competing risks survival model, incorporating one CT imaging feature and baseline information, showed significantly improved performance for predicting the onset of severe disease (mean time-dependent area under the curve (AUC) = 0.880). Multiple predispositional factors can be utilized to assess the risk of severe COVID-19 onset at an early stage. Multivariate survival models can reasonably estimate progression risk based on early-stage CT images, which can be easily misjudged in qualitative analysis.

2. Materials and methods

2.1. Study cohort

The cohort included all adult inpatients admitted to the Third People's Hospital of Shenzhen (Shenzhen, China) between January 11, 2020, and February 29, 2020. The hospital is a designated center for the admission of COVID-19 patients from other hospitals in Shenzhen—a city in southern China characterized by a large proportion of migrant workers. All patients had tested positive for SARS-CoV-2 infection at other local hospitals, and these results were confirmed at the Shenzhen CDC during patient admission. Upon hospitalization, all patients were confirmed to have SARS-CoV-2 infection but at different stages of the disease. As of March 8, 2020,

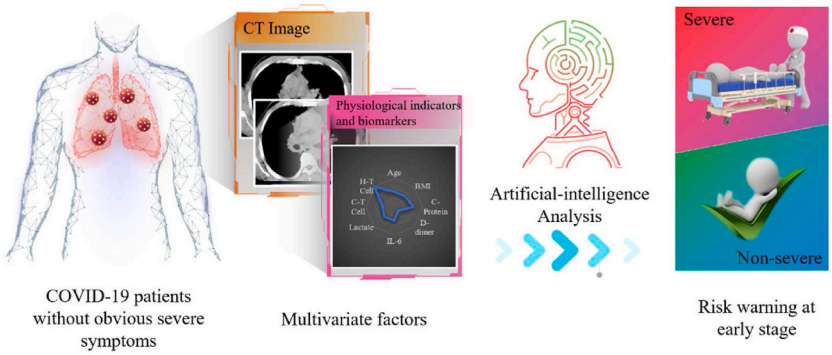


Fig. 1. Severity-onset prediction of COVID-19 via an artificial-intelligence analysis of multivariate factors.

290 (85.8%) of 338 patients had recovered from COVID-19 and were discharged from the hospital. In general, patients received antiviral therapy upon admission, and if severe progression occurred, they received ventilation support and/or glucocorticoid-related therapies according to the WHO guidelines [13]. This study was approved by the Ethics Committee of the Third People's Hospital of Shenzhen (2020-099). Written informed consent was obtained from all participating patients.

2.2. Statistical analysis

Continuous and categorical variables are presented as mean (standard deviation (SD)) and number (%). The significance of between-group differences was evaluated using a one-way analysis of variance (ANOVA), chi-square test, or Fisher's exact test. We checked the normality assumption of the ANOVA test using both quantile–quantile plotting and Shapiro–Wilk testing of the residuals generated by the ANOVA model. For the quantile–quantile plot, the normality assumption was considered not violated if most of the data points followed the straight $y = x$ line. For the Shapiro–Wilk test, a p value greater than 0.05 suggested that the residuals were normally distributed. The Wilcoxon rank-sum test was used when data did not meet the assumption of normality. Multiple testing was corrected using the Benjamini–Hochberg procedure [14] to control the false discovery rate and to obtain adjusted p values. An adjusted p value smaller than 0.01 was considered statistically significant. Multivariate logistic regression was used to test the significance of the interaction effects, whenever possible. All statistical analyses were performed using R software (version 3.6.1).

2.3. Computer-aided CT image processing and feature extraction

We used "simple ITK" [15] to read the original CT sectional images and then reconstructed them into 3D images. Six hundred images were used for model training and ninety-seven images were used for model testing. The input images were transformed into 3D images that were subsequently computed as voxels. The voxels were resized to $1\text{mm} \times 1\text{mm} \times 1\text{mm}$ to ensure equal distances between neighboring voxels in all directions. Otsu's method, which was fine-tuned using a Laplacian filter, was applied to segment the lung tissue area in a 3D context. A morphological operation was used to polish the segmentation results. Part of the lung airways was identified and eliminated as previously described [16]. The features were computed on the original image or on images that were further processed using a 3D Laplacian of Gaussian filter with different sigma values (1, 2, 3, 4, and 5 mm). For each derived image, we quantified 18 first-order features, 24 Gy-level co-occurrence matrix features, and 16 Gy-level size zone matrix features, resulting in 348 features for each patient. These features are listed in Table S4. The image analyses were performed using Python software (version 3.7.6) with the "scikit-image" [17] and "Py Radiomics" [18] packages. To extract features from the original CT images, we designed an artificial-intelligence model to analyze the multivariate factors. The model is based on convolutional neural networks. First, we used three convolutional neural network layers for feature extraction and a rectified linear unit as the activation function. Two residual neural network blocks were used for the analysis of multivariate factors. We used the softmax function as the output layer to obtain the results. All the frameworks were constructed using PyTorch and an NVIDIA A6000 GPU.

2.4. Time-to-event data analysis and survival modeling

We let T be the time of event occurrence and D be the event type or cause. The competing risks were handled by the cause-specific hazard function, which is the hazard of developing an event of a given cause in the presence of competing events, and can be written as

$$\lambda_k(t) = \lim_{\Delta t \longrightarrow 0} \frac{Pr\{t \le T \le t + \Delta t, D = k | \ T \ge t\}}{\Delta t}$$

The cumulative incidence function of cause k, $I_k(t)$, interpreted as the cumulative probability $\Pr(T \le t, D = k)$ of developing an event of cause k before time t, can be expressed as

$$I_k(t) = \int_0^t \lambda_k(t) S(s) ds$$

where the survival function can be defined as $S(t) \ = exp(-\int_0^t \lambda_k(t) ds)$. To analyze the underlying risk pattern of the time-to-event data, we used the Fine-Gray model [19,20], in which patients with an event of another cause remain in the risk set for parameter estimation.

To incorporate the covariate $\boldsymbol{Z}$, we applied the strategy of imposing the proportional hazard assumption on cause-specific hazards [21], which can be defined as

$$\lambda_k(t|\boldsymbol{Z}) = \lambda_{k,0}(t) \exp(\boldsymbol{\beta_k Z})$$

where $\lambda_{k,0}(t)$ is the baseline hazard for cause k, and $\boldsymbol{\beta_k}$ are the covariate coefficients for cause k. The features were selected by applying the Lasso [22] or elastic net [23] shrinkage methods to the cause-specific approach mentioned above. The optimal shrinkage parameter was selected based on three-fold cross-validation results. We computed the risk prediction score by multiplying the selected features and their coefficients and further assessed the significance of the score using the Fine-Gray model, where the proportional hazard assumption is imposed on the cumulative incidence function. We used time-dependent receiver operating characteristic [24] and bootstrapping of 632+ prediction error curves [25] for prediction performance evaluation. Because of the limited sample size, we only report the internal cross-validation results. The analyses and modeling were performed using R software (version 3.6.1) and with R packages "survival," "cmprsk," "pec," and "risk Regression." The algorithms are publicly available through the website medicinedata.

cn/covid-19-prognosis.

3. Results

3.1. Disease process and clinical outcomes

All 338 adult patients enrolled in this study were treated at Shenzhen Third People's Hospital. The typical disease progression of a patient is shown in Fig. 2. After hospital admission, the patient either progresses to a severe condition or recovers from pneumonia without severe progression. Among the 76 (31.9%) patients who developed severe disease, 18 (5.3%) progressed to critical condition. As of March 8, 2020, 3 (0.9%) patients had died, and 45 (13.3%) remained hospitalized. All the other patients were discharged from the hospital with in a cured state. The mean duration from symptom onset to hospital admission was 5.1 days, and the mean time from admission to the onset of the severe disease was 3.7 days (Table S1).

3.2. Characteristics of the study cohort

To evaluate progression-related features, we classified the patients into severe and non-severe groups based on disease severity. The time intervals between symptom onset and hospital admission were not significantly different between the two groups ($p = 0.264$). Patient age was bimodally distributed, with one mode located at approximately 35 years and the other at 60 years (Fig. 3 A). As shown in Table S2, the mean age of patients in the severe group was significantly higher than that of patients in the non-severe group (58.7 and 46.1, respectively; Fig. 3 B). Moreover, the severe group included patients that were more overweight (Fig. 3C). Most patients (79%) had a travel history to the Wuhan region, the epicenter of the initial outbreak, within 14 days of symptom onset. Sex was not strongly associated with the disease severity. Blood type was not associated with the severity outcome.

The patients' medical histories were not significantly worse in the severe group; however, coexisting hypertension or diabetes was significant. The most common symptoms upon admission were fever (60.7%) and cough (52.4%). The severe group had a significantly higher rate of fever than the non-severe group.

3.3. Laboratory testing results

The testing was performed on blood samples that were collected immediately after hospital admission. The severity classifications of patients are summarized in Table S3. The severe group had a significantly lower number of platelets and lymphocytes and increased levels of coagulation function indicators, such as fibrinogen, *D*-dimer and activated partial thromboplastin time. In the blood biochemistry tests, the biomarkers that were significantly increased in the severe group compared with the non-severe group included lactate dehydrogenase, myoglobin, aspartate transaminase, mitochondrial aspartate transaminase, creatine kinase myocardial band, troponin I, N-terminal brain natriuretic peptide, γ-glutamyl transpeptidase, and alpha-hydroxybutyrate dehydrogenase; however, the biomarkers that were decreased included albumin and prealbumin. For infection-related biomarkers, we observed significant increases in C-reactive protein, interleukin-6, procalcitonin, and erythrocyte sedimentation rates in the severe group, all of which had mean levels beyond the upper limit of the normal reference range. We also observed abnormal blood oxygen levels, as indicated by the PaO_2/FiO_2 ratio, and increased levels of kidney function indicators, including glomerular filtration rate, cystatin C and β2-microglobulin. Lactic acid tests in patients during the early stage were not associated with severity outcome.

To investigate the pattern of dynamic variation in the representative biomarkers and immune T cells, we plotted these biomarker concentrations and T cell counts against time since symptom onset, stratified by severity classification. Compared with the non-severe group, the severe group had increased levels of C-reactive protein, which tended to converge at the late stage during hospitalization for the two groups (Fig. 3 D). The D-dimer levels in the severe group gradually increased 7 days after symptom onset and were significantly higher than those in the non-severe group (Fig. 3 E). Interleukin-6 in the severe group gradually decreased during the first 10 days after symptom onset, which was equivalent to that in the non-severe group, and then continued to increase (Fig. 3 F). The lactate dehydrogenase levels in the severe group increased in the first 10 days and decreased in the next 10 days. Subsequently, they increased

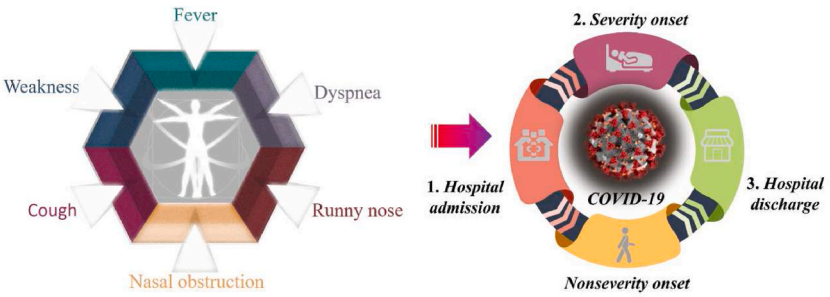
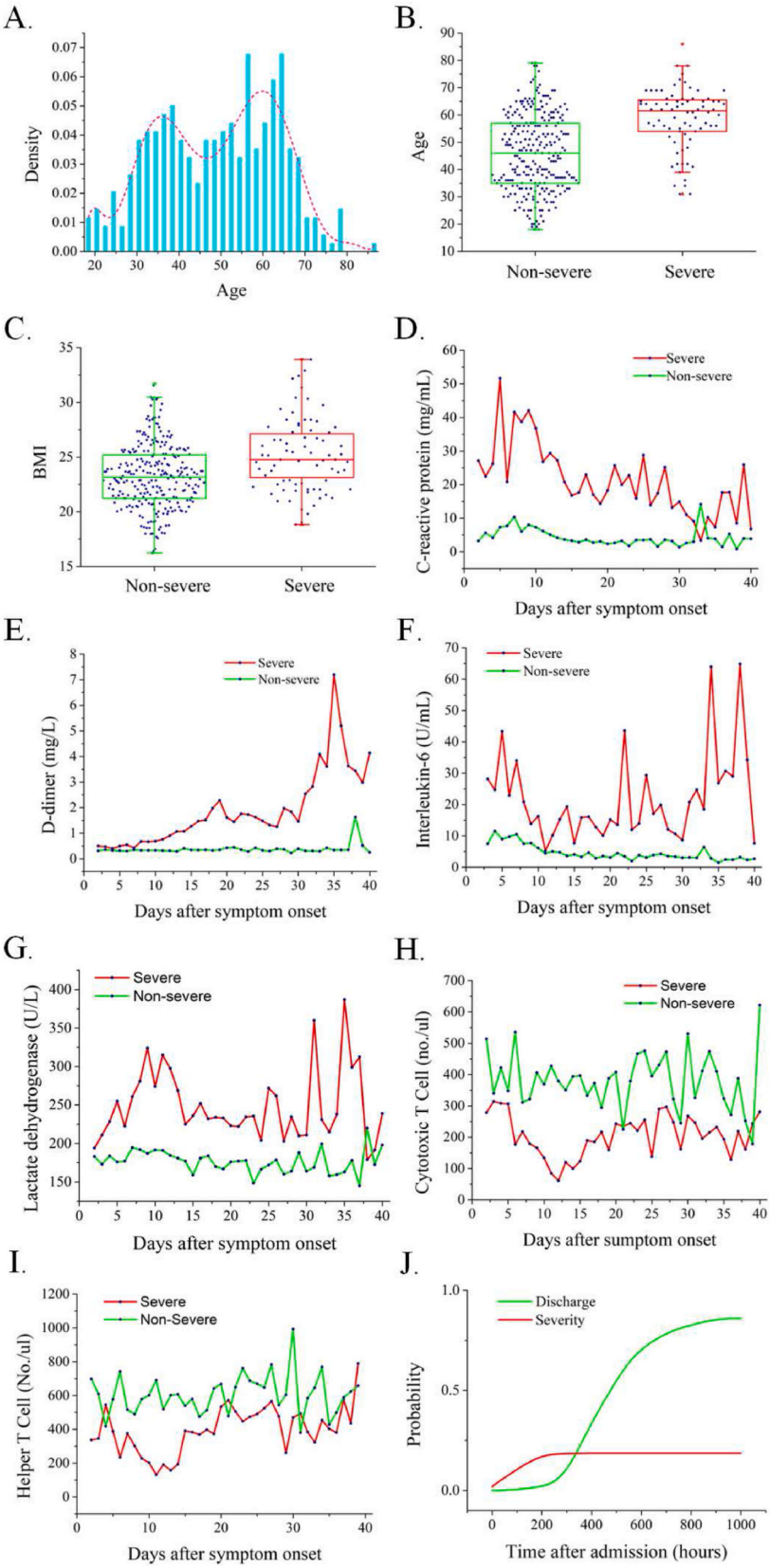


Fig. 2. Diagram of the disease process for patients with COVID-19. The hospital discharge of patients with no disease progression after admission (1 to 3) poses a competing risk for those patients who are progressing to a severe condition (1 to 2).



(caption on next page)

Fig. 3. Analysis of the representative features that were significantly associated with the severe group. A, the age distribution of the study cohort, which is overlaid with the kernel density estimates (solid curve); B and C, boxplot summary of the age and BMI as stratified by the severe and non-severe group; D-I, the levels of the individual laboratory biomarkers (C-reactive protein, D-dimer, interleukin-6, lactate dehydrogenase) and immune T cells (Cytotoxic T cell, Helper T cell) are plotted along with the time after the symptom onset; I, cumulative probability for the patients in the study cohort of developing a severe condition or being discharged from hospital after hospital admission. Data from multiple patients on the same day were combined, and only the median value (solid dots) is shown.

again (Fig. 3 G). In addition, the severe group showed decreased numbers of cytotoxic and helper T cells in the first 10 days after symptom onset. This number increased within the next 10 days and plateaued thereafter (Fig. 3 H–I).

3.4. Pattern of disease progression

We performed a competing risk analysis of the time-to-event data to assess the risk pattern of disease progression during hospitalization. The cumulative probability (the cumulative incidence function) of the onset of severe disease continued to increase

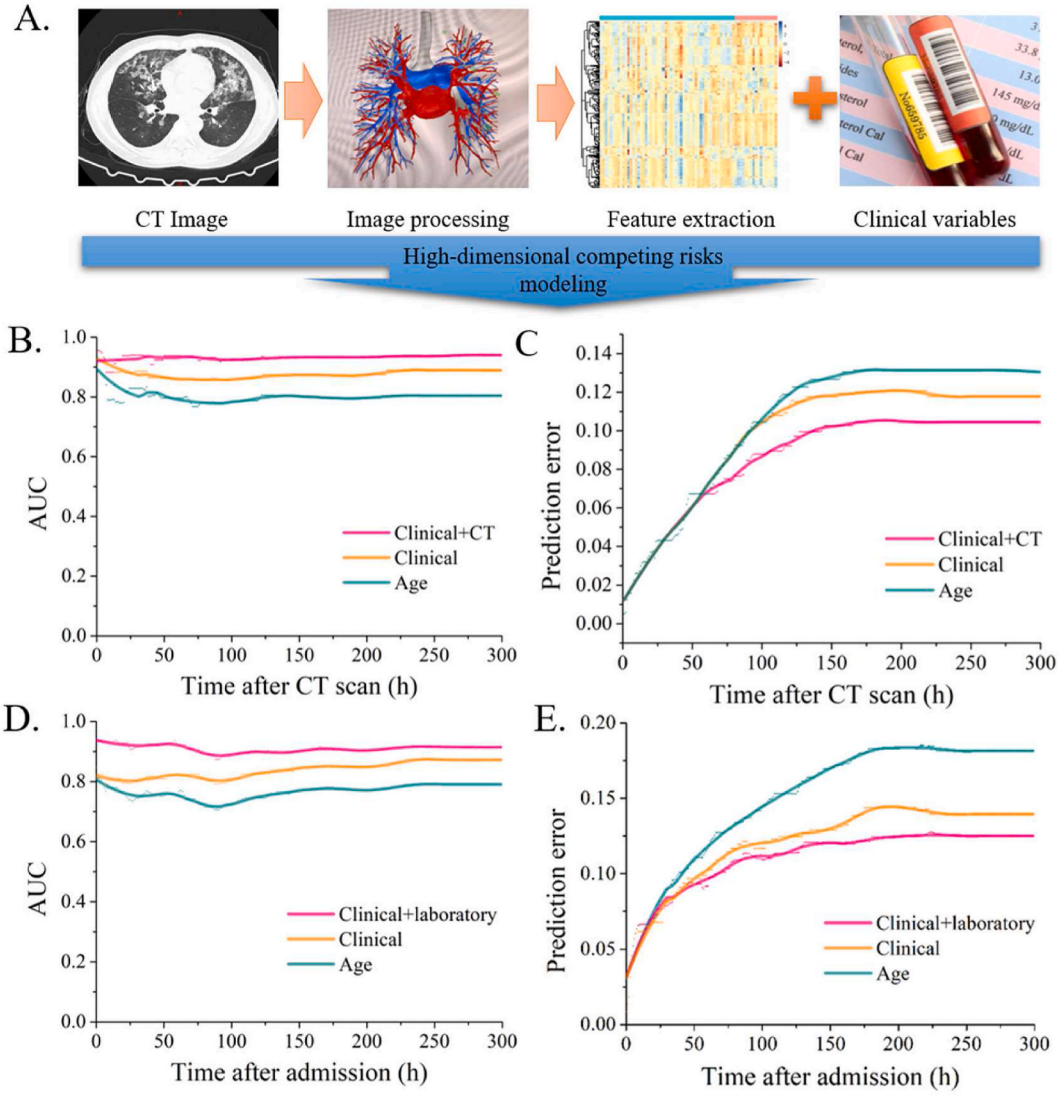


Fig. 4. Workflow of the development of the CT-based risk prediction model. A, Heatmap analysis of the quantitative features that were extracted from the CT images, stratified by the progression outcome; B, the evaluation of the model prediction performance by the time-dependent ROC method; C, evaluation of the model prediction performance by the time-dependent prediction error method; D. the prediction performance evaluation results of the laboratory testing-based risk assessment model by the time-dependent ROC (left) and E. by the time-dependent prediction error (right).

immediately after hospital admission but reached its change point on day 10 (Fig. 3 J). In contrast, the cumulative incidence of the competing risk event, which was discharge because of recovery, was much lower during the high-risk period of onset of severe disease. However, the incidence increased dramatically from approximately day 12 to 29.

3.5. Risk prediction of progression

To personalize the time-dependent risk assessment, we incorporated the CT imaging data with baseline information for statistical modeling (Fig. 4). A total of 348 quantitative features were extracted from the 3D reconstructed chest CT images generated upon admission. Cluster analysis revealed a subset of features that could distinguish the severe group from the non-severe group (Fig. 4 A). We then computed the same features from the earliest CT scans generated after admission for each patient and trained a risk prediction score using high-dimensional survival modeling. The model integrating the CT and baseline variables significantly outperformed the univariate and multivariate models using only the baseline information (Fig. 4 B and C). The best model achieved a mean time-dependent AUC of 0.880 (SD = 0.011) and a mean prediction error of 0.079 (SD = 0.024). We also developed a model that integrated laboratory biomarkers tested within one day of admission. This model had a mean AUC of 0.884 (SD = 0.049) and mean prediction error of 0.103 (SD = 0.031) (Fig. 4 D and E).

We present two cases to demonstrate the use of the CT image-based risk assessment tool (Fig. 5). Patients in cases I and II were similar in terms of age and BMI (Table 1). The patient in case I had unilateral ground-glass opacities on CT images, whereas the patient in case II had more obvious bilateral opacities (Fig. 5 A). However, the patient in case I, rather than that in case II, progressed to a severe condition. Follow-up CT scans also showed deteriorating conditions for the patient in case I (Fig. 5C). Consistently, our model predicted the cumulative probabilities of developing severe disease for case I within the next 1, 3, and 5 days as 0.032, 0.073, and 0.121, respectively, compared with 0.001, 0.003, and 0.005 for case II, respectively (Fig. 5 B). The presence of fever symptoms on admission of the patient in case I added discriminative power to the model.

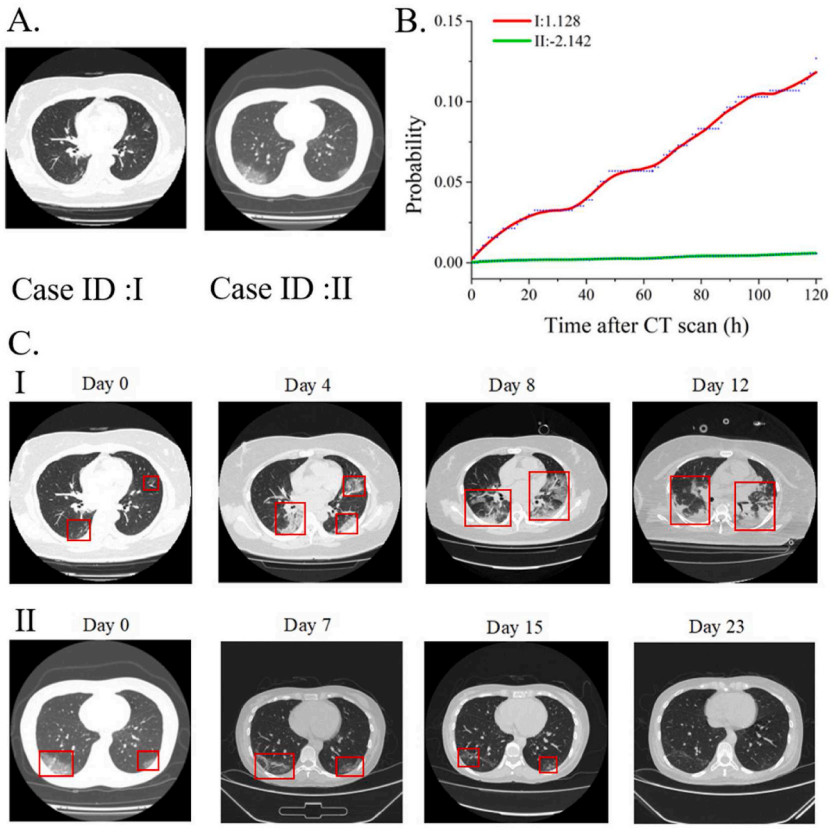


Fig. 5. Illustration of applying the risk prediction model in a clinical setting. A, Baseline information about the example cases. The time interval indicates the time between the CT scan and severity onset if there was severe progression or the time between admission and hospital discharge if there was no such progression; B, representative CT scan of the example cases; C, Follow-up CT scans for case I and case II as presented. The title of each figure is the days since the first scan; D, Risk prediction results of the example cases. The prediction score is shown in the figure legend.

Table 1
Clinical characteristics of the example cases.

Case ID	Age	BMI	Fever at admission	Sever progression	Time interval (h)
I	36	21	Yes	Yes	167
II	37	19	No	No	574

4. Discussion

We delineated the characteristics of a retrospective cohort of 338 adult patients from a single center in Shenzhen City, China, and developed a noninvasive method to evaluate the risk of severe COVID-19 onset. Independent predisposing factors for disease progression include old age, high BMI, fever, and coexisting hypertension or diabetes [26,27]. However, using age as a single prognostic factor could lead to erroneous results since young patients are not necessarily progression-free. In contrast to previous studies [28–30], the severe group in this cohort had a significantly higher proportion of patients with fever symptoms at admission (82.9% vs. 54.2%). These findings benefit the risk assessment analysis, as we showed that the model that combined these indicators can substantially improve prediction performance compared with the model that only contains univariate predictors (mean time-dependent AUC = 0.824 versus 0.751).

SARS-CoV-2 enters the host cells through the angiotensin-converting enzyme 2 receptor [31], whose expression can be enhanced by the use of hypertension medicine, such as angiotensin-converting enzyme inhibitors or angiotensin II type-I receptor blockers [32]. Whether such medications are a causal factor for a higher risk of COVID-19 progression is currently under investigation [33]. In our study, we observed that the presence of underlying hypertension as a comorbidity was significantly related to severe progression; however, we did not find an association between medications and severe outcomes among patients with hypertension.

COVID-19 pneumonia is a multistate disease with clinically relevant intermediate endpoints such as the onset of severe disease. Most survival data analyses set the onset as the primary endpoint and censored patient recovery or hospital discharge. However, when competing risks of severe disease onset are present, these analytical methods introduce bias. In this study, the risk of severe progression was assessed without considering that these competing risks would be overestimated because patients who would never progress (those who were discharged from the hospital without progression) were treated as if they could progress. The extent of such bias and its adjustment by competing risk modeling have been evaluated in clinical virological and oncological research [34–37]. We incorporated high-dimensional variable selection techniques into competing risk modeling so that quantitative image features could be extensively evaluated according to their contribution to risk prediction. Our evaluation results showed that incorporating CT images into the model can significantly improve the prediction performance compared with models based only on demographic and clinical information (mean time-dependent AUC = 0.880 versus 0.824). Such an improvement was achieved with additional image features, suggesting the importance of using multimodal data in risk analysis.

The laboratory testing results in this study showed that, at the time of admission, patients in the severe group already had signs of functional impairment in organs such as the liver (lactate dehydrogenase and prealbumin), heart (multiple types of myocardial enzymes including troponin I, N-terminal brain natriuretic peptide, creatine kinase myocardial band, mitochondrial aspartate transaminase, and aspartate transaminase), and kidneys (glomerular filtration rate, cystatin C, and β2-microglobulin). Abnormalities in blood oxygen levels were also observed, such as abnormalities in thePaO_2/FiO_2, which was used to determine the onset of severe COVID-19 in this study, and this value was significantly, although not ideally, different between the two groups upon admission. Consistent with previous studies [9,29,30,38,39], the severe group in this cohort had, at an early stage, a substantial increase in inflammatory factors, such as C-reactive protein and interleukin-6, and coagulation dysfunction. Incorporating laboratory biomarkers tested at an early stage can also significantly improve the risk prediction performance compared with the best model that does not include them (mean AUC = 0.884 versus 0.813).

Our study has a few possible limitations. First, laboratory testing data were not integrated into the CT image-based model because blood samples were not collected along with imaging in the center. A more complicated statistical model is required to account for the errors caused by inherent time differences in the input data. Second, the severity may depend on other factors, such as treatment, viral load, or genetic factors. In the future, the model can include such additional covariates. However, the availability of well-processed data on these factors is low. We have shown that the current model can achieve reasonable prediction performance even without considering these factors. Third, because of the limited sample size of patients who progressed to severe disease, we only evaluated model prediction performance using a cross-validation analysis. Further validation must be conducted using external datasets. Fourth, we restricted our analysis to the severe group until the onset of severe disease. Factors related to recovery from severe conditions have not been evaluated. Finally, this study did not include young patients (age <18 years).

Ethics statement

The study was approved by the Ethical Committee of the Third People's Hospital of Shenzhen (2020-099).

Data sharing statement

All data are available in the main text or the supplementary materials. All materials are available from the co-responding author upon reasonable request.

Author contribution statement

Yuhui Liao: Lijiao Zeng: Conceived and designed the experiments; Wrote the paper.
Yu Fu: Conceived and designed the experiments; Performed the experiments; Wrote the paper.
Pilai Huang: Mingfeng Liao: Jialu Li: Qinlang Shi: Performed the experiments.
Mingxia Zhang: Performed the experiments; Contributed reagents, materials, analysis tools or data.
Zhaohua Xia: Xinzhong Ning: Jiu Mo: Ziyuan Zhou: Zigang Li: Analyzed and interpreted the data.
Jing Yuan: Lifei Wang: Qing He: Qikang Wu: Lei Liu: Contributed reagents, materials, analysis tools or data.
Kun Qiao: Conceived and designed the experiments; Wrote the paper.

Data availability statement

Data will be made available on request.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by the National Key R&D Program of China (2021YFC2302200), the National Natural Science Foundation of China (81972019, 21904145, 82002253), China Postdoctoral Science Foundation (2021M691428), Special Fund of Foshan Summit Plan (2020B019, 2020B012, 2020A015), Training project of National Science Foundation for Outstanding/Excellent Young Scholars of Southern Medical University (C620PF0217), Regional Joint Fund of Natural Science Foundation of Guangdong Province (2020A1515110529), Guangdong Basic and Applied Basic Research Foundation (2020A1515010754), the Fundamental Research Funds for the Central Universities (2019MS134), Chen Jingyu Team of Sanming Project of Medicine in Shenzhen (SZSM201812058), Shenzhen Science and Technological Foundation (JSGG20210901145200001), Guangdong Medical science foundation(A2021413). Natural Science Foundation of China grants (22107045). Hospital Fund of Chinese Academy of Medical Sciences Cancer Hospital Shenzhen Hospital(E010221005).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2023.e18764.

References

[1] S.B. Kotwal, N. Orekondey, G.P. Saradadevi, N. Priyadarshini, N.V. Puppala, M. Bhushan, S. Motamarry, R. Kumar, G. Mohannath, R.J. Dey, Multidimensional futuristic approaches to address the pandemics beyond COVID-19, Heliyon 9 (6) (2023), e17148.
[2] M. Ciaccio, L. Agnello, Biochemical biomarkers alterations in coronavirus disease 2019 (COVID-19), Diagnosis (Berlin, Germany) 7 (4) (2020) 365–372.
[3] G. Xue, X. Gan, Z. Wu, D. Xie, Y. Xiong, L. Hua, B. Zhou, N. Zhou, J. Xiang, J. Li, Novel serological biomarkers for inflammation in predicting disease severity in patients with COVID-19, Int. Immunopharm. 89 (Pt A) (2020), 107065.
[4] S.K. Ray, S. Mukherjee, The spectrum of biochemical alterations with molecular and serological biomarkers in the diagnosis of COVID-19: searching for novel one to identify disease earlier with better prognosis and drug discovery, Recent advances in anti-infective drug discovery 16 (3) (2021) 179–181.
[5] E.T. Qian, C.L. Gatto, O. Amusina, M.L. Dear, W. Hiser, R. Buie, S. Kripalani, F.E. Harrell Jr., R.E. Freundlich, Y. Gao, W. Gong, C. Hennessy, J. Grooms, M. Mattingly, S.K. Bellam, J. Burke, A. Zakaria, E.E. Vasilevskis, F.T.t. Billings, J.M. Pulley, G.R. Bernard, C.J. Lindsell, T.W. Rice, Assessment of awake prone positioning in hospitalized adults with COVID-19: a nonrandomized controlled trial, JAMA Intern. Med. 182 (6) (2022) 612–621.
[6] S.R. Knight, A. Ho, R. Pius, I. Buchan, G. Carson, T.M. Drake, J. Dunning, C.J. Fairfield, C. Gamble, C.A. Green, R. Gupta, S. Halpin, H.E. Hardwick, K.A. Holden, P.W. Horby, C. Jackson, K.A. McLean, L. Merson, J.S. Nguyen-Van-Tam, L. Norman, M. Noursadeghi, P.L. Olliaro, M.G. Pritchard, C.D. Russell, C.A. Shaw, A. Sheikh, T. Solomon, C. Sudlow, O.V. Swann, L.C. Turtle, P.J. Openshaw, J.K. Baillie, M.G. Semple, A.B. Docherty, E.M. Harrison, Risk stratification of patients admitted to hospital with covid-19 using the ISARIC WHO Clinical Characterisation Protocol: development and validation of the 4C Mortality Score, BMJ 370 (2020) m3339.
[7] C. Huang, Y. Wang, X. Li, L. Ren, B. Cao, Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China, Lancet 395 (10223) (2020) 497–506.
[8] S. Fink, F. Ruoff, A. Stahl, M. Becker, P. Kaiser, B. Traenkle, D. Junker, F. Weise, N. Ruetalo, S. Hörber, A. Peter, A. Nelde, J. Walz, G. Krause, A. Baillot, K. Schenke-Layland, T.O. Joos, U. Rothbauer, N. Schneiderhan-Marra, M. Schindler, M.F. Templin, Multiplexed serum antibody screening platform using virus extracts from endemic coronaviridae and SARS-CoV-2, ACS Infect. Dis. 7 (6) (2021) 1596–1606.
[9] F. Zhou, T. Yu, R. Du, G. Fan, B. Cao, Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study, Lancet 395 (10229) (2020) 1054–1062.
[10] H. Shi, X. Han, N. Jiang, Y. Cao, O. Alwalid, J. Gu, Y. Fan, C. Zheng, Radiological findings from 81 patients with COVID-19 pneumonia in Wuhan, China: a descriptive study, Lancet Infect. Dis. 20 (4) (2020) 425–434.
[11] X. Wang, H. Zhang, H. Du, R. Ma, Y. Nan, T. Zhang, Risk factors for COVID-19 in patients with hypertension, Can. J. Infect Dis. Med. Microbiol. 2021 (2021) 1–9.
[12] R.W. Alberca, L.M. Oliveira, A. Branco, N.Z. Pereira, M.N. Sato, Obesity as a risk factor for COVID-19: an overview, Crit. Rev. Food Sci. Nutr. 61 (13) (2021) 2262–2276.

[13] O. World Health, Clinical Management of Severe Acute Respiratory Infection when Novel Coronavirus (2019-nCoV) Infection Is Suspected: Interim Guidance, 28 January 2020, World Health Organization, Geneva, 2020.
[14] Y. Benjamini, Y. Hochberg, Controlling the false discovery rate: a practical and powerful approach to multiple testing, J. Roy. Stat. Soc. B: Methodological 57 (1) (1995) 289–300.
[15] B.C. Lowekamp, D.T. Chen, L. Ibáez, D. Blezek, The design of simpleITK, Front. Neuroinf. 7 (45) (2013) 45.
[16] J. Wang, F. Li, Q. Li, Automated segmentation of lungs with severe interstitial lung disease in CT, Med. Phys. 36 (10) (2009) 4592–4599.
[17] S.V.D. Walt, J.L. Schnberger, J. Nunez-Iglesias, F. Boulogne, S.I, Contributors, scikit-image: image processing in Python, PeerJ 2 (2) (2014) e453.
[18] J.J.M. van Griethuysen, A. Fedorov, C. Parmar, A. Hosny, N. Aucoin, V. Narayan, R.G.H. Beets-Tan, J.C. Fillion-Robin, S. Pieper, H. Aerts, Computational Radiomics system to decode the radiographic phenotype, Cancer Res. 77 (21) (2017) e104–e107.
[19] J.P. Fine, R.J. Gray, A proportional hazards model for the subdistribution of a competing risk, Publ. Am. Stat. Assoc. 94 (446) (1999) 496–509.
[20] R.B. Geskus, Cause-specific cumulative incidence estimation and the fine and gray model under both left truncation and right censoring, Biometrics 67 (1) (2011) 39–49.
[21] D.J. Holt, Competing risk analyses with special reference to matched pair experiments, Biometrika 65 (1) (1978) 159–165.
[22] R. Tibshirani, Regression shrinkage and selection via the lasso, J. Roy. Stat. Soc. B 58 (1) (1996) 267–288.
[23] H. Zou, T. Hastie, Regularization and variable selection via the elastic net, J. Roy. Stat. Soc. B 67 (2) (2005) 301–320.
[24] P.J. Heagerty, T. Lumley, M.S. Pepe, Time-dependent ROC curves for censored survival data and a diagnostic marker, Biometrics 56 (2) (2000) 337–344.
[25] H. Binder, A. Allignol, M. Schumacher, J. Beyersmann, Boosting for high-dimensional time-to-event data with competing risks, Bioinformatics 25 (7) (2009) 890–896.
[26] S. Lin, X. Deng, I. Ryan, K. Zhang, W. Zhang, E. Oghaghare, D.B. Gayle, B. Shaw, COVID-19 symptoms and deaths among healthcare workers, United States, Emerg. Infect. Dis. 28 (8) (2022) 1624–1641.
[27] Y.K. Ko, H. Murayama, L. Yamasaki, R. Kinoshita, M. Suzuki, H. Nishiura, Age-dependent effects of COVID-19 vaccine and of healthcare burden on COVID-19 deaths, tokyo, Japan, Emerg. Infect. Dis. 28 (9) (2022) 1777–1784.
[28] C. Wu, X. Chen, Y. Cai, X. Zhou, S. Xu, H. Huang, L. Zhang, X. Zhou, C. Du, Y. Zhang, Risk factors associated with acute respiratory distress syndrome and death in patients with coronavirus disease 2019 pneumonia in wuhan, China, JAMA Intern. Med. 180 (7) (2020) 934–943.
[29] D. Wang, B. Hu, C. Hu, F. Zhu, X. Liu, J. Zhang, B. Wang, H. Xiang, Z. Cheng, Y. Xiong, Y. Zhao, Y. Li, X. Wang, Z. Peng, Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus–infected pneumonia in wuhan, China, JAMA 323 (11) (2020) 1061–1069.
[30] C. Huang, Y. Wang, X. Li, L. Ren, J. Zhao, Y. Hu, L. Zhang, G. Fan, J. Xu, X. Gu, Z. Cheng, T. Yu, J. Xia, Y. Wei, W. Wu, X. Xie, W. Yin, H. Li, M. Liu, Y. Xiao, H. Gao, L. Guo, J. Xie, G. Wang, R. Jiang, Z. Gao, Q. Jin, J. Wang, B. Cao, Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China, Lancet 395 (10223) (2020) 497–506.
[31] Y. Wan, J. Shang, R. Graham, R.S. Baric, F. Li, Receptor recognition by the novel coronavirus from Wuhan: an analysis based on decade-long structural studies of SARS Coronavirus, J. Virol. 94 (7) (2020) 10–1128.
[32] X.C. Li, J. Zhang, J.L. Zhuo, The vasoprotective axes of the renin-angiotensin system: physiological relevance and therapeutic implications in cardiovascular, hypertensive and kidney diseases, Pharmacol. Res. 125 (2017) 21–38.
[33] L. Fang, G. Karakiulakis, M. Roth, Are patients with hypertension and diabetes mellitus at increased risk for COVID-19 infection, Lancet Respir. Med. 8 (4) (2020) e21.
[34] O.M. Rueda, S.J. Sammut, J.A. Seoane, S.F. Chin, J.L. Caswell-Jin, M. Callari, R. Batra, B. Pereira, A. Bruna, H.R. Ali, E. Provenzano, B. Liu, M. Parisien, C. Gillett, S. McKinney, A.R. Green, L. Murphy, A. Purushotham, I.O. Ellis, P.D. Pharoah, C. Rueda, S. Aparicio, C. Caldas, C. Curtis, Dynamics of breast-cancer relapse reveal late-recurring ER-positive genomic subgroups, Nature 567 (7748) (2019) 399–404.
[35] H. Putter, M. Fiocco, R.B. Geskus, Tutorial in biostatistics: competing risks and multi-state models, Stat. Med. 26 (11) (2007) 2389–2430.
[36] R.P. van Rij, A.M. de Roda Husman, M. Brouwer, J. Goudsmit, R.A. Coutinho, H. Schuitemaker, Role of CCR2 genotype in the clinical course of syncytium-inducing (SI) or non-SI human immunodeficiency virus type 1 infection and in the time to conversion to SI virus variants, J. Infect. Dis. 178 (6) (1998) 1806–1811.
[37] R.L. Prentice, J.D. Kalbfleisch, A.V. Peterson Jr., N. Flournoy, V.T. Farewell, N.E. Breslow, The analysis of failure times in the presence of competing risks, Biometrics 34 (4) (1978) 541–554.
[38] X. Yang, Y. Yu, J. Xu, H. Shu, J. Xia, H. Liu, Y. Wu, L. Zhang, Z. Yu, M. Fang, T. Yu, Y. Wang, S. Pan, X. Zou, S. Yuan, Y. Shang, Clinical course and outcomes of critically ill patients with SARS-CoV-2 pneumonia in Wuhan, China: a single-centered, retrospective, observational study, Lancet Respir. Med. 8 (5) (2020) 475–481.
[39] Z. Shang, S.Y. Chan, W.J. Liu, P. Li, W. Huang, Recent insights into emerging coronavirus: SARS-CoV-2, ACS Infect. Dis. 7 (6) (2021) 1369–1388.