Reviewer A

Comment 1: My only major criticism is can authors incorporate more data (which is now available) especially different geographic areas; especially data from US and including some data on viral loads as subsets has to be included.

Reply 1: Thank you for your valuable comments, which have provided important guidance for my future research. In this study, the proportion of the severe group was relatively lower, and there was an imbalance between the sizes of the severe and non-severe group. However, more data we could incorporate were mainly imported cases from abroad, and severe cases were rare. Therefore, in the case of relatively low proportion of severe cases, we modified the statistics analysis. Base on the optimal clinic radiological features selected by the LASSO logistic regression, a nomogram model for prediction was established by five-fold cross validation. More data, especially from different geographic areas, are needed to validate the robustness of the model so as to further improve efficiency. This is one of the limitations of this study and also the direction of future research.

Changes in the text: We have added to the discussion of this limitation (see Page 16, line 308-313).

Comment 2: It is really important to include group of people's data who recovered from disease and who exposed – those will serve important internal controls to validate the models and its accuracy in predicting disease severity.

Reply 2: Thank you for your valuable comments. Since it was the early stage of the outbreak, our focus was on whether patients would become severe illness. As you said,

it is also very important to observe the development and prognosis of the disease. Thank you for providing important guidance for my future research.

Reviewer B

Comment: Please add the predictor of severe COVID-19 from others' studies in the discussion, and recently several CT-based AI models has been reported, it would be better the authors also included them in the discussion part.

Reply: Thank you for your encouraging remarks and valuable comments. Again, some time passed, more researches have been reported on the disease. Current researches suggest elevated levels of acute phase reactants are early predictors of high disease severity. Our study suggests $CD4^+$ T cells, lactate dehydrogenase and C-reactive protein could be used as predictors for the severity of COVID-19, consistent with current studies. Recent studies have reported promising results of the use of AI combined with CT imaging for diagnosing COVID-19. By studying the latest literatures, we can deepen our understandings of this field.

Changes in the text: We have modified our text as advised in the discussion part (see Page 13, line 248-258 and Page 15, line 283-290).

Reviewer C

Comment 1: Validity assessment of models wasn't conducted using internal and external validation in this study.

Reply 1: Thank you for your valuable comments. In our study, the proportion of the severe group was relatively lower, and there was an imbalance between the sizes of the severe and non-severe group. However, more data we could incorporate were mainly imported cases from abroad, and severe cases were rare. Therefore, in the case of relatively low proportion of severe cases, we modified the statistics analysis. Base on the optimal clinicoradiological features selected by the LASSO logistic regression, a nomogram model for prediction was established by five-fold cross validation. In the

future, we need to include more data, especially from different geographic regions, to verify the robustness of the model so as to further improve efficiency.

Changes in the text: We have added to the discussion of this limitation (see Page 16, line 308-313).

Comment 2: The disease severity in the COVID 19 patients was classified into mild, medium, severe and critical clinical types according to the diagnostic criteria from the Chinese Clinical Guidance for COVID-19 Pneumonia Diagnosis and Treatment. I think the approach in the guidelines is much more convenient than using a model that includes a CT test to distinguish the severity of the disease.

Reply 2: Thank you for your valuable comments. The Chinese Clinical Guidance are of practical value in distinguishing between severe and non-severe cases. The goals of our model are both to identify severe illness and to judge whether patients will become critically ill. However, our model needs further validation and improvement.

Comment 3: The model is not visualized, such as drawing up a nomogram and risk scoring model.

Reply 3: Thank you for your valuable comments. Adding this part makes the study more complete and our model more practical.

Changes in the text: We have modified our text as advised (see Page 9-10, line 170-174, line 181-182, Page 12, line 226-230, Figure 6 and Figure 7).