Peer Review File

Article information: https://dx.doi.org/10.21037/atm-21-4126

Reviewer #A

To develop a model with 96 patients to predict survival seems completely underpowered and irrelevant to me. As I understand from the abstract, the model has discriminative power of 80% at 9 days after hospital admission. It seems completely logic that one can predict survival of a viral infection after 9 days of hospitalisation and I do not understand the clinical need. Therefore, cannot make time to review the complete manuscript. After reading the abstract, I would not accept this paper for publication.

Reply: Thank you for this important comment. We agree that the multivariate joint model with 96 patients to predict survival may underpowered. Therefore, we have considered it as a limitation in this manuscript and mentioned that the findings need to be generalized with caution to the other studies. All the same, the multivariate joint model takes longitudinal data into consideration, which provides more information on disease progression and improves the prediction accuracy effectively. Besides, it can characterize the time-to-event process, obtain dynamic predictions at the individual level, and describe time-varying associations between the longitudinal biomarkers and the event. We can predict the daily survival probabilities of each patient after admission, which facilitates the identification of critical time points and can alert the clinician when to apply patient-tailored therapies to decrease the case-fatality rate.

Changes in the text: We have modified in the Abstract Section, which states *"Results: Random Forest selected a set of important biomarkers including C-reactive protein (CRP), blood urea nitrogen (BUN), procalcitonin (PCT), base excess (BE), Lymphocyte count (LYMPH), white blood cell (WBC), and creatine phosphokinase (CPK) with the lowest classification error in the feature selection phase. The multivariate joint model was used to describe the effect of biomarkers and characterize the dynamic progression of the event. Combined with the covariates, the joint model demonstrated a good performance in discriminating the survived and deceased patients within a fixed time window of 3 days. And the AUC were stable at about 0.75 during hospitalization.*
Conclusion: Our study established a novel model that can identify important indicators associated with the prognostic outcome of H7N9, characterize the time-to-event process and predict the daily survival probabilities after admission at the individual level." (See page 3, lines 47-57).

Reviewer #B

Zhang et al. described their findings in establishment of the dynamic model for disease prognosis prediction among H7N9 infected patients. Their results indicated that combined with the covariates, the multivariate joint model demonstrated a promising performance in predicting the prognostic outcome. Although this study offers important information regrading establishment of a novel prediction model for H7N9 disease outcome and prognosis, however, there are still many points needed further clarification.

Major comments:

Comment 1: *Lines 65-67, the authors indicated that that biomarkers, such as oxygenation index, neutrophil percentage, C-reactive protein, and white blood cell, play an essential role in the H7N9 progression and are independent predictors of the survival outcome. However, there are many biomarkers reported to be correlated with H7N9 disease progression, such as Angiotensin II in plasma levels, HLA-DR levels of CD14+ cells, certain cytokines, etc. Please try to update the information.*

Reply 1: Thank you for the suggestion. We have updated the information in the revised manuscript.

Changes in the text: We added the biomarkers reported to be correlated with H7N9 disease progression in the Introduction Section, which states *"Previous studies have found that biomarkers, such as oxygenation index, neutrophil percentage, C-reactive protein, white blood cell, cytokines, Angiotensin II in plasma levels, HLA-DR levels of CD14+ cells, and so on play an essential role in the H7N9 progression and are independent predictors of the survival outcome."* (See page 4, lines 68-71).

Comment 2: *There was a lack of ethic statement, because collection and analyses of patients' data still needed informed consent and approved by Institutional Review Boards.*

Reply 2: Thank you for this important comment. We have added the ethic statement in the Footnote Section. (See page 15, lines 325-330)

Comment 3: *In the Method section, how did the patients confirm as H7N9 infection? which kind of diagnostic assays were used?*

Reply 3: Thank you for this important comment. The patients were diagnosed by real-time fluorescent quantitative PCR assay, and we have added the information in the revised manuscript.

Changes in the text: We added the information in the Methods Section, which states *"A confirmed case of H7N9 can be diagnosed by real-time fluorescent quantitative PCR assay when nucleic acid of H7N9 virus is detected in upper respiratory tract specimens (pharyngeal swabs) or deep respiratory tract specimens (sputum or bronchoalveolar lavage fluid) with severe bilateral pneumonia, leukopenia, and lymphocytopenia, and the H7N9 virus strain can be isolated from some specimens."* (See page 6, lines 103-108).

Comment 4: *What kinds of laboratory indicators were selected for calculation? the rationale?*

Reply 4: Thank you. It was a retrospective observational study, and the laboratory indicators we selected were the necessary clinical data of patients during hospitalization, such as blood routine and blood biochemistry.

Comment 5: *Lines 158-159, The recruited patients showed that higher percentage of men were infected. Please explain this difference, any bias occurrence when patient recruitment? In addition, some comorbidities, signs and symptoms as well as Laboratory findings showed statistical significance when compared the survivor with the death. How to properly select these factors for disease dynamic or prognosis prediction after H7N9 infection?*

Reply 5: Thank you for this important comment. All H7N9-infected patients in Suzhou, Wuxi, Huai'an, and Taizhou in Jiangsu Province, China between January 2016 and May 2017 were recruited, and previous studies also showed that men with a higher proportion of infections (1, 2). Besides, it was a retrospective observational study, and as such, subject selection bias was avoidable.

The important biomarkers were selected through the sliding windows sequential forward feature selection method (SWSFS), which was data-driven. Specifically, the set of biomarkers having the lowest model error were screened out as candidate prognostic factors for further analysis. The details of this method were described in the

Methods section (See page 7, lines 129-137).

Comment 6: *The potential influence factors which may lead to inaccuracy prediction should be discussed.*

Reply 6: Thank you, we appreciate these suggestions. We do agree that bias and confounders may lead to inaccuracy prediction. And we have adjusted the covariates as much as possible. However, we can't fully take unmeasured confounders and possible bias into consideration due to data collection. Therefore, we considered it as a limitation in this manuscript.

Changes in the text: We have modified the Discussion Section of the manuscript accordingly, which states *"Second, unmeasured confounders and possible bias may affect the prediction accuracy of the model."* (See page 13, lines 285-286).

Comment 7: *The Results and Conclusion of the Abstract should mention which biomarkers or indicators found in this study were properly used in dynamic model established regarding H7N9 disease progresses.*

Reply 7: Thanks for this suggestion. We have mentioned the biomarkers associated with the prognostic outcome of H7N9 in the revised manuscript.

Changes in the text: We have modified the Abstract Section of the manuscript accordingly, which states *"Random Forest selected a set of important biomarkers including C-reactive protein (CRP), blood urea nitrogen (BUN), procalcitonin (PCT), base excess (BE), Lymphocyte count (LYMPH), white blood cell (WBC), and creatine phosphokinase (CPK) with the lowest classification error in the feature selection phase."* (See page 3, lines 47-50).

1. European Centre for Disease Prevention and Control. Human infection with avian influenza A(H7N9) virus –fourth update. Rapid Risk Assessment 2015, Stockholm.
2. Centers for Disease Control and Prevention (CDC). Emergence of avian influenza A(H7N9) virus causing severe human illness - China, February-April 2013. MMWR Morb Mortal Wkly Rep 2013; 62:366-371.