Peer Review File
Article information: https://dx.doi.org/10.21037/jtd-24-416

Reviewer A

Predicting frailty risk is an important goal in COPD care. This observational study addressed factors determining risk for frailty in patients with COPD. Strengths of the study include machine learning approaches for analysis.

Major comments:

Comment1: Abstract: The Abstract, whilst quite well-written, does not clearly provide the definition of frailty used in the analysis, which should be included. Also, the sample size studied is not included in the Abstract.

Reply 1: We sincerely thank the reviewers for their careful review. We have added the definition of frailty. "Frailty is a medical syndrome caused by multiple factors and characterized by reduced physical strength, endurance and physiological functioning, leading to increased susceptibility to dependency and/or death." The details are set out in the "Abstract" "of the article (see page 1, lines 30-32). We have added the sample size of the study. The details can be found in the “Methods” section of the article (see Page 4, line 120-121).

Comment2: External validation: As the authors mentioned in their Discussion, no external validation had been undertaken, which limits the generalizability of the results of this study.

Reply 2: We think this is a very good suggestion. This is indeed a weakness of our study and we will endeavour to validate with externally collected clinical data in future studies to improve the reliability of the study.

Comment3: Frailty measures: A composite group of frailty measures was used. Was this a validated collection of variables used in previous publications, or was this used only in this study, and if so, how were these variables prioritized and selected (over others).

Reply 3: Thank you for your kind words. A comprehensive set of frailty measures was used. It has not been used in previous studies for the time being, and was only used in this study. These variables were prioritized and selected based on a combination of clinical guidelines on COPD, the occurrence of frailty and relevant research.

Comment4: Patients: How were the participants recruited and enrolled for the CHARLS database, and how were the current participants selected for this study? Further details are required, to assess for any selection bias.

Reply 4: Thank you so much for your careful review. The China Health and Retirement Longitudinal Study (CHARLS) is an ongoing longitudinal survey designed to investigate the social, economic, and health status of middle-aged and older adults aged 45 and over in China. The baseline survey was conducted in 2011 with a total of 17,708 participants who are followed up every two years, and there are now four periods of data available, 2011 (wave 1), 2013 (wave 2), 2015 (wave 3), and 2018 (wave 4). For a more detailed description of the database, please refer to the official website or Cohort Profile: The China Health and Retirement Longitudinal Study (CHARLS).

Comment5: Accuracy: Are the results of the XGBoost model considered accurate enough for clinical use, based on the effect sizes seen?

Reply 5: The results of the XGBoost model are accurate and can be used clinically. But more patients are needed to validate the results.

Minor comments:

Typographical changes:

Page 1: 'prediction model for frailty inpatients with COPD' – change to 'in patients'; otherwise this could be confused with 'inpatients' admitted to hospital

Reply: We feel great thanks for your professional review work on our article. We have changed on page 1: "Predictive Modelling of frailty inpatients with COPD" to "Predictive Modelling of frailty in patients with COPD". The details can be seen in the "Abstract" of the article (see Page 2, line 35).

Reviewer B

This original article presents a sophisticated, well-executed study on the development and validation of machine learning-based prediction models for frailty occurrence in patients with Chronic Obstructive Pulmonary Disease (COPD). Using a comprehensive dataset from the 2018 China Health and Retirement Longitudinal Study and deploying an array of machine learning techniques, the study identifies key predictors of frailty and evaluates the predictive power of different models, with the XGBoost model showing exceptional performance.

Strengths:

1. **Comprehensive Methodology**: The study meticulously combines 34 indicators, including behavioral, health status, mental health, and sociodemographic variables, offering a holistic approach to frailty prediction in COPD patients.

2. **Innovative Use of Machine Learning**: By employing seven different machine learning techniques and demonstrating the superior performance of the XGBoost model, the study is at the forefront of applying advanced computational methods to clinical research.
3. **Practical Application**: The creation of an online predictive risk modeling website, complemented by SHAP interpretations, enhances the study's practical value, allowing for easy access and interpretation by healthcare professionals.
4. **Significant Findings**: Identifying key predictors such as depression, smoking, and gender provides actionable insights for healthcare providers to identify at-risk patients effectively.
Limitations:
Comment1: **Generalizability**: The study's reliance on data from China might limit its applicability to other populations without further validation.
Reply 1: Thank you for your review. We will try to combine the large foreign databases in our future studies to minimize the discrepancies caused by different populations. The "Discussion" section of the study in the article presents a detailed explanatory account of our efforts to remedy the shortcomings (see page 15, lines 476-493).

Comment2: Missing Factors**: The exclusion of certain potential predictive factors such as lung function and BMI could impact the model's comprehensiveness and accuracy.
Reply 2: Sincerely, thank you for your advice. As this study used data from The China Health and Retirement Longitudinal Study (CHARLS) database, some of the clinical indicators were not collected, which is a shortcoming of our study. The "Discussion" section of the study in the article presents a detailed explanatory account of our efforts to remedy the shortcomings (see page 15, lines 476-493).

Comment3: **Retrospective Nature**: Being a retrospective study based on existing database information might limit the ability to capture all relevant predictors and dynamics of frailty development over time.
Reply 3: Thank you for your comment. This is indeed our shortcoming as well, and we need to conduct prospective studies in future research. The "Discussion" section of the study in the article presents a detailed explanatory account of our efforts to remedy the shortcomings (see page 15, lines 476-493).

Score: 9/10
Rationale for the Score: This article contributes significantly to the field of predictive health analytics for COPD patients, demonstrating a high level of innovation,

methodological rigor, and practical utility. The minor limitations noted, primarily related to the study's scope and data source, do not significantly detract from the overall quality and impact of the research.

Reply 4: Thank you for your comments.