Research Proposal
Predicting the Onset of Invasive Mechanical Ventilation in Patients with COPD: A Machine Learning Approach

Yehoshua Friedman, Research Methods, JCT Data Mining 2022-2023

Introduction:

Chronic Obstructive Pulmonary Disease (COPD) is a leading cause of morbidity and mortality worldwide. The disease is characterized by persistent respiratory symptoms and airflow limitation, often leading to severe hypoxemia. Invasive Mechanical Ventilation (IMV) is a life-saving intervention for COPD patients experiencing acute respiratory failure. However, the decision to initiate IMV is complex and multifaceted, requiring careful consideration of the patient's current health status, disease progression, and potential for recovery. Delayed initiation of IMV can lead to poor patient outcomes, including increased mortality. Therefore, there is a pressing need for robust, reliable tools to predict the necessity for IMV in COPD patients.

This research proposal aims to address this need by developing a machine learning model to predict the onset of IMV in COPD patients. The proposed model will be based on a two-step approach, similar to the one used in a recent study predicting IMV necessity in critically ill COVID-19 patients. The first step involves training a predictive model on a large dataset of non-COPD critically ill hypoxemic patients. The second step applies transfer learning to adapt the model to a smaller COPD cohort. Chronic Obstructive Pulmonary Disease (COPD) is a progressive lung disease characterized by increasing breathlessness, frequent coughing, wheezing, and tightness in the chest. It is a major cause of morbidity and mortality worldwide, and its management poses significant challenges to healthcare systems. One of the most critical aspects of managing severe COPD cases is the decision to initiate invasive mechanical ventilation (IMV). This decision is often complex and can significantly impact patient outcomes.

Research Motivations

A study by Bendavid et al. (2022) demonstrated the potential of machine learning models in predicting the need for IMV in critically ill patients suffering from COVID-19. The researchers developed a two-step model, first training a machine learning predictive model on a large dataset of non-COVID-19 critically ill hypoxemic patients, and then applying transfer learning to adapt the model to a smaller COVID-19 cohort. The model was able to accurately predict the need for IMV 6, 12, 18, or 24 hours in advance, potentially aiding the decision-making process in patients with hypoxemic respiratory failure. The success of the model developed by Bendavid et al. (2022) in predicting the need for IMV in COVID-19 patients provides a promising foundation for the application of similar models in other respiratory diseases, such as COPD. Given the complexity and critical nature of the decision to initiate IMV in COPD patients, a machine learning model that can accurately predict the need for IMV could significantly improve patient outcomes and optimize the use of healthcare resources.

Furthermore, the ability to predict the need for IMV in advance could potentially allow for earlier interventions that may prevent the progression of the disease to a stage where IMV is required. This could not only improve patient outcomes but also reduce the burden on healthcare systems.

Preliminary Data:

The feasibility and rationale for this research proposal are grounded in preliminary experiments and findings from previous studies, particularly the study on predicting the need for IMV in critically ill COVID-19 patients using a two-step machine learning model.

In the referenced study, a machine learning model was trained on a large dataset of non-COVID-19 critically ill hypoxemic patients from the United States (MIMIC-III). The model was then adapted to a smaller COVID-19 cohort using transfer learning techniques. The model successfully predicted the need for IMV 6, 12, 18, or 24 hours in advance in both the general ICU population and COVID-19 patients. The model demonstrated good predictive power, with an AUC of 0.83 on a shortened set of features, excluding the clinician's settings, and an AUC of 0.91 when the clinician settings were included.

These preliminary findings provide strong evidence for the feasibility of developing a similar machine learning model for predicting the onset of IMV in COPD patients. The successful application of the two-step model in the COVID-19 study suggests that a similar approach could be effective for COPD patients. The high AUC values achieved in the COVID-19 study indicate that the model was able to accurately predict the need for IMV, supporting the rationale for this research proposal. Additionally, the use of transfer learning in the COVID-19 study demonstrates the potential for adapting the model to different patient populations. This is particularly relevant for this research proposal, as the model will be initially trained on a non-COPD patient cohort and then adapted to a COPD patient cohort.

Furthermore, a study conducted in 2022 used machine learning to predict successful extubation in patients receiving mechanical ventilation (Frontiers in Medicine, 2022). The findings from this study underscore the potential of machine learning models in predicting outcomes related to mechanical ventilation. Additionally, a study published in PLOS ONE in 2021 used machine learning methods to predict mechanical ventilation and mortality in patients with COVID-19 (PLOS ONE, 2021). This study further supports the feasibility of our research proposal, as it demonstrates the potential of machine learning models in predicting critical outcomes in patients with respiratory diseases.

These preliminary findings provide strong evidence for the feasibility of developing a similar machine learning model for predicting the onset of IMV in COPD patients. The successful application of machine learning models in these studies suggests that a similar approach could be effective for COPD patients. The high predictive accuracy achieved in these studies supports the rationale for this research proposal.

Research Question and Hypothesis:

Research Question: Can a machine learning model, trained and validated on patient data, accurately predict the onset of Invasive Mechanical Ventilation (IMV) in patients with Chronic Obstructive Pulmonary Disease (COPD)?

Hypothesis: Based on the results of the study by Bendavid et al. (2022) on COVID-19 patients, we hypothesize that a machine learning model can be developed and trained on COPD patient data to accurately predict the onset of IMV. The model in their study achieved an area under the receiver operating characteristic curve (AUC-ROC) of 0.83 on a shortened set of features, excluding the clinician's settings, and an AUC-ROC of 0.91 when the clinician settings were included, indicating a high level of accuracy in predicting the need for IMV in critically ill patients. Therefore, we hypothesize that a similar model developed for COPD patients will achieve comparable AUC-ROC values. This prediction capability could potentially aid in the decision-making process for initiating IMV in COPD patients, thereby improving patient outcomes and optimizing the use of healthcare resources.

Research Methodology:

The research will be conducted in several stages:

1) Data Collection and Preprocessing:
   The first step will involve collecting and preprocessing data from two distinct patient cohorts: a large cohort of non-COPD critically ill hypoxemic patients and a smaller cohort of COPD patients. The data will be collected from electronic health records (EHRs) and will include both static variables (e.g., age, sex, comorbidities) and dynamic measurements (e.g., vital signs, laboratory results). The preprocessing stage will involve cleaning the data, handling missing values, and normalizing the data to ensure it is suitable for machine learning. This stage will also involve feature engineering, where new features may be created based on the existing data to enhance the predictive power of the model.

2) Model Development:
   The second stage will involve developing the machine learning model. The model will be trained on the large dataset of non-COPD critically ill hypoxemic patients. Several machine learning algorithms will be tested, including logistic regression, random forest, and gradient boosting. The model's performance will be evaluated using cross-validation, and the best-performing model will be selected for further refinement.

3) Transfer Learning:
   The third stage will involve applying transfer learning techniques to adapt the model to the smaller COPD cohort. This will involve fine-tuning the model parameters to optimize its performance for COPD patients. The performance of the adapted model will be evaluated using cross-validation.

4) Model Validation:
   The fourth stage will involve validating the performance of the adapted model in an independent cohort of COPD patients. The model's predictive accuracy, sensitivity, specificity, and area under the receiver operating characteristic curve (AUC-ROC) will be evaluated.

Conclusion:

The successful completion of this research will not only contribute to the existing body of knowledge on COPD management but also has the potential to significantly improve patient outcomes. By enabling early intervention, the model can help reduce the risk of complications associated with delayed IMV initiation. Furthermore, by predicting the need for IMV, the model can assist in optimal resource allocation in intensive care units, thereby enhancing the efficiency of healthcare delivery.

This research proposal represents a significant step forward in the application of machine learning in healthcare, particularly in the management of chronic diseases like COPD. The methodologies and insights gained from this research could potentially be applied to other chronic diseases, paving the way for more personalized and effective patient care.

References:

1) ResearchGate. (2020). A novel machine learning model to predict respiratory failure and invasive mechanical ventilation in critically ill patients suffering from COVID-19. Retrieved from https://www.researchgate.net/publication/361479210_A_novel_machine_learning_model_to_predict_respiratory_failure_and_invasive_mechanical_ventilation_in_critically_ill_patients_suffering_from_COVID-19
2) Frontiers in Medicine. (2023). Developing a machine-learning model for real-time prediction of successful extubation in mechanically ventilated patients using time-series ventilator-derived parameters. Retrieved from https://www.frontiersin.org/articles/10.3389/fmed.2023.1167445/full
3) Frontiers in Medicine. (2022). Machine learning for predicting successful extubation in patients receiving mechanical ventilation. Retrieved from https://www.frontiersin.org/articles/10.3389/fmed.2022.961252/full
4) PLOS ONE. (2021). Machine learning methods to predict mechanical ventilation and mortality in patients with COVID-19. Retrieved from https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0249285