**Study Protocol**

**Development and validation of patient-level prediction models for disease progression amongst adult patients admitted to hospital with pneumonia: a rapid network study to inform the management of COVID-19**

​

**Created by:**

Henrik John, Erasmus MC, Rotterdam, The Netherlands  
Aniek Markus, Erasmus MC, Rotterdam, The Netherlands

Tom M. Seinen, Erasmus MC, Rotterdam, The Netherlands

Ross D. Williams, Erasmus MC, Rotterdam, The Netherlands

Cynthia Yang, Erasmus MC, Rotterdam, The Netherlands

Jan Kors, PhD, Erasmus MC, Rotterdam, The Netherlands

Jenna Reps, PhD, Janssen Research and Development

Peter Rijnbeek, PhD, Erasmus MC, Rotterdam, The Netherlands

**Prepared on:** 28 March 2020

**Acknowledgement:** The analysis is performed in the context of the European Health Data and Evidence Network (EHDEN) project ([www.ehden.eu](http://www.ehden.eu/)) in close collaboration with the Observational Health Sciences and Informatics collaborative (OHDSI, [http://ohdsi.org](http://ohdsi.org/)).

The authors declare the following disclosures: Jenna Reps, PhD is employee of Janssen Research and Development.

**Table of Contents**

[1. List of Abbreviations 3](#_Toc36400421)

[2. Executive Summary 3](#_Toc36400422)

[3. Rationale & Background 3](#_Toc36400423)

[4. Objective 5](#_Toc36400424)

[5. Methods 6](#_Toc36400425)

[5.1. Study Design 6](#_Toc36400426)

[5.2. Data Source(s) 6](#_Toc36400427)

[5.3. Study Populations 7](#_Toc36400428)

[5.4. Statistical Analysis Method(s) 9](#_Toc36400429)

[5.5. Quality Control 10](#_Toc36400430)

[5.6. Tools 10](#_Toc36400431)

[6. Diagnostics 10](#_Toc36400432)

[7. Data Analysis Plan 11](#_Toc36400433)

[7.1. Algorithm Settings 11](#_Toc36400434)

[7.2. Covariate Settings 11](#_Toc36400435)

[7.3. Model Development & Evaluation 16](#_Toc36400436)

[7.4. Analysis Execution Settings 16](#_Toc36400437)

[8. Strengths & Limitations 16](#_Toc36400438)

[9. Protection of Human Subjects 16](#_Toc36400439)

[10. Plans for Disseminating & Communicating Study Results 17](#_Toc36400440)

[11. Tables & Figures 17](#_Toc36400441)

[11.1. Incidence Rate of Target & Outcome 17](#_Toc36400442)

[12. Appendices 17](#_Toc36400443)

[12.1. Study Generation Version Information 17](#_Toc36400444)

[13. References 17](#_Toc36400445)

# List of Abbreviations

|  |  |
| --- | --- |
| Abbreviation | Phrase |
| AUROC | Area Under the Receiver Operating Characteristic Curve |
| CDM | Common Data Model |
| O | Outcome Cohort |
| OHDSI | Observational Health Data Sciences & Informatics |
| OMOP | Observational Medical Outcomes Partnership |
| T | Target Cohort |
| TAR | Time at Risk |

# Executive Summary

The objective of this study is to develop and validate patient-level prediction models for adult patients that are being hospitalized with pneumonia, to predict two outcomes: 1) hospitalizations with pneumonia, acute respiratory distress syndrome, sepsis, acute kidney injury, requiring intensive services or resulting in death, and 2) persons who die. Both with a time at risk of 30 days from hospital admission. These two prediction models will be implemented using Lasso Logistic Regression.

# Rationale & Background

General description of COVID-19

The Corona Virus Disease 2019 (COVID-19), which started in late 2019 as an epidemic in Wuhan, Hubei Province, China, has been declared a pandemic and a public health emergency of international concern by the World Health Organization (WHO) in January 2020 (1). The growing number of infections by COVID-19 has resulted in an unprecedented pressure on healthcare systems worldwide, and a large number of casualties at a global scale. Diagnosis of COVID-19 currently relies on the detection of SARS-CoV-2 nucleic acid (2); no cure or vaccine is available yet. Common symptoms presented by patients include fever, cough, and dyspnea, signaling the onset of pneumonia (3). Although the majority of people have uncomplicated or mild illness (81%), some will develop severe illness requiring hospitalization and oxygen support (14%) or intensive care unit treatment (5%) (4).

Problem definition

There is a lack of data and evidence on the factors associated with disease severity and/or mortality of patients diagnosed with COVID-19. While the number of infected patients continues to increase globally, the pressure on healthcare systems increases as well. The rapid increase in severely ill patients has resulted in an immense shortage of resources and available ICU beds (5-6). Due to this scarcity, knowledge on which patients are at high risk and should therefore require close monitoring is valuable. This knowledge may also be used to project future demand of ICU care. Early reports on COVID-19 cases have shown that it takes on average about 5 days from having symptoms to developing severe illness, with pneumonia as the most common diagnosis. In the same vein, it generally takes about 10 days from having symptoms to develop critical illness, which is defined as acute respiratory distress syndrome (ARDS) or sepsis with acute organ dysfunction (3). Correctly identifying which patients will benefit most from close monitoring will ensure these patients have the best chance of receiving optimal care at the right time and enhance their chances of recovery. This may prevent further progress to complications associated with critical illness and consequently reduce the number of ICU admissions.

Study aims

The objective of this study is to inform the management of adult patients who are hospitalized with COVID-19 by developing and validating a patient-level prediction model. In particular, we aim to use medical history information prior to admission to identify which patients are at high risk of developing complications associated with critical illness and/or mortality. However, due to the rapid onset of the COVID-19 pandemic, a current barrier to producing a patient-level prediction model for patients with COVID-19 is the still limited numbers of patients that are readily available to study. Since pneumonia is the most common serious manifestation of COVID-19 infection, we have chosen to use pneumonia as a proxy for COVID-19, although pneumonia caused by COVID-19 may be particularly severe. After developing a model on patients who are admitted to hospital with pneumonia, we will then validate in the COVID-19 patient datasets as they become available. If the models are shown to be transportable then this will increase the speed at which they can be disseminated and as such have a greater impact on the attempt to control the most negative impacts of the pandemic.

Clinical use case

This prediction model can be used by health care providers to identify upon hospital admission which patients are at high risk of developing complications associated with critical illness and/or mortality and will require close monitoring. It can also be used by health care administrators to determine whether their hospital has sufficient health care resources to account for the impending number of predicted patients requiring intensive services. In this way, prognosis of patients may be improved while resources may be used more optimally.

Description of previous literature

We reviewed previous literature on predicting risk of mortality and complications including admission to ICU, mechanical ventilation and ARDS in pneumonia patients. Since COVID-19 is infectious and contagious in the general population, we focused on literature on pneumonia contracted by a person outside of the healthcare system, referred to as community-acquired pneumonia (CAP). Many studies examined patients admitted to hospital with CAP and most of these studies aimed to predict outcomes within 30 days of hospital admission (7-11). Outcomes of interest focused on include mortality, admission to ICU and other adverse outcomes while hospitalized.

The Pneumonia Severity Index (PSI) score for CAP estimates the mortality for adult patients with CAP, using 20 different variables (7). Other more simpler scores for pneumonia severity, such as CURB-65 and A-DROP, also estimate mortality of CAP to help determine inpatient vs. outpatient treatment and have similar performance as PSI (8, 12). Other scores, such as SMART-COP predicts the need for intensive respiratory or vasopressor support in CAP (13). Recently a new score, MuLBSTA, has been published but was not externally validated (14). Candidate predictors that have been considered vary between the reviewed studies, but many scores rely primarily on demographics, simple physiological measurements, and present comorbidities. Almost all the mentioned scores use a time as risk of 30 days and have an AUROC of around 0.75 and 0.85 in both the initial articles and subsequent validation studies.

Additionally, we reviewed recent COVID-19 prediction modelling studies. Some recent Chinese studies constructed prediction models based on cohorts of COVID-19 patients. One study aimed to construct and validate a model for early identification of severe cases among hospitalized non-severe COVID-19 patients (15). A nomogram was developed that contained one clinical (age) and six serological indicators and was externally validated. The nomogram showed excellent discrimination (AUC 0.91 internal, 0.86 external validation) and calibration. Another study developed a simple mortality risk index based on COVID-19 patients (16). This index, called ACP index, contains only two variables, age and C-reactive protein. The ACP index was used to stratify the 12-day mortality risk of COVID-19 patients in three grades of disease severity. No standard performance measures were reported. Although these studies provide useful insight into possibilities of managing COVID-19 patients, most studies are not able to externally validate their prediction models or are constructed on very limited data. We add to the literature by developing our models using data from previous flu/viral outbreaks, studying sufferers of previous infections and pneumoniae. We speculate that any learning from previous viral outbreaks will be to some degree applicable to the current COVID-19 pandemic and will collaborate with colleagues from heavily affected areas of the world to externally validate our findings as soon as their data is available for research.

How the study is performed

In this study we aim to develop and validate a patient-level prediction model to identify, based on the medical history information prior to hospital admission with COVID-19, which patients are at high risk of developing complications associated with critical illness and/or mortality. We will develop a patient-level prediction model using the Observational Health Data Sciences and Informatics (OHDSI) Patient-Level Prediction framework. The OHDSI collaboration is a network of researchers working towards a common goal of standardizations and best practice frameworks for analyzing observational data in healthcare. The OHDSI collaboration relies on researchers mapping their datasets into the Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM). This along with the Patient-Level Prediction framework allows for rapid model development and validation following accepted best practices (17). We will implement the patient-level prediction models across databases across the OHDSI collaborator network to externally validate the models and evaluate their transportability across the world.

# Objective

The objective is to develop and validate patient-level prediction models for disease progression amongst adult patients admitted to hospital with pneumonia:

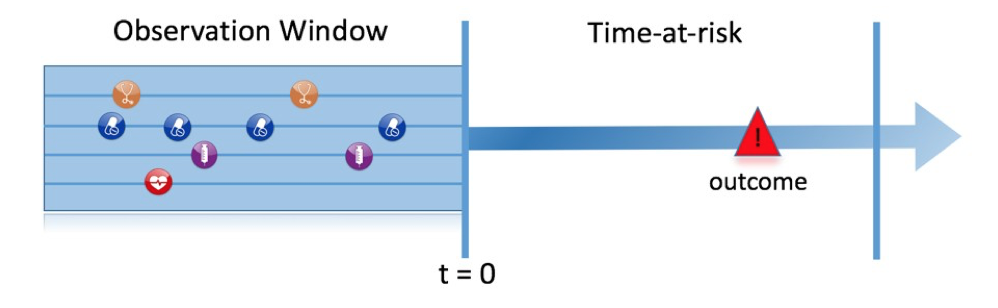
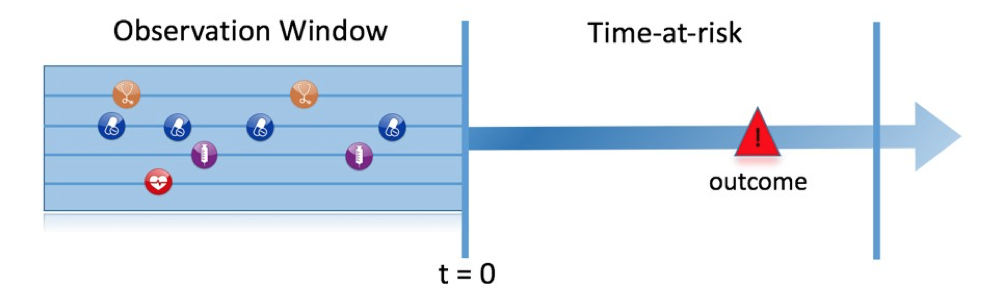
* To predict the 30-day risk of developing complications associated with critical illness and/or mortality amongst adult patients admitted to hospital with pneumonia.

# Methods

## Study Design

This study will follow a retrospective, observational, patient-level prediction design. We define 'retrospective' to mean the study will be conducted using data already collected prior to the start of the study. We define 'observational' to mean there is no intervention or treatment assignment imposed by the study. We define 'patient-level prediction' as a modelling process wherein an outcome is predicted within a time at risk relative to the target cohort start and/or end date. Prediction is performed using a set of covariates derived using data prior to the start of the target cohort.

Figure 1 illustrates the prediction problem we will address. Among a population at risk, we aim to predict which patients at a defined moment in time (t = 0) will experience some outcome during a time-at-risk (TAR). Prediction is done using only information about the patients in an observation window prior to that moment in time.



***Figure 1: The prediction problem***

We follow the PROGRESS best practice recommendations for model development and the TRIPOD guidance for transparent reporting of the model results (18, 19).

## Data Source(s)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Source Full Name** | **Country Code** | **Data Provenance** | **Source Short Name** | **Patient Count** | **History** | **Patient Type** | **Data collection** |
| Optum® de-identified Electronic  Health Record Dataset | US | EMR | Optum EHR - EMR, US | 96m | 2006- | EHR / Privately Insured | Optum© de-identified Electronic Health Record Dataset represents Humedica’s Electronic Health Record data a medical records database. The medical record data includes clinical information, inclusive of prescriptions as prescribed and administered, lab results, vital signs, body measurements, diagnoses, procedures, and information derived from clinical Notes using Natural Language Processing (NLP). |

## Study Populations

### Target Cohort(s) [T]

|  |  |  |
| --- | --- | --- |
| Cohort ID | Cohort Name | Description |
| 5891 | [COVID19 ID29 V1] Hospitalizations with pneumonia, age>=18 | TBD |

### Outcome Cohorts(s) [O]

|  |  |  |
| --- | --- | --- |
| Cohort ID | Cohort Name | Description |
| 5889 | [COVID19 ID27 V1] Hospitalizations with pneumonia or ARDS or sepsis or AKI requiring intensive services or resulting in death in 30d | TBD |
| 5890 | [COVID19 ID28 v1] persons who die | TBD |

Full descriptions:

The JSON files describing for all the outcome cohorts are available at: <https://github.com/ohdsi-studies/Covid19PredictionStudies/tree/master/SevereInHospitalizedPatients/inst/cohorts>

In order to convert these to a human readable form, import the JSON into a new cohort definition in any instance of ATLAS and reload.

### Time at Risk

The table below describes the Time at Risk (TAR) window start and end for each of the analyses that are executed.

|  |
| --- |
| Time at Risk |
| [Time at Risk Settings #1] Risk Window Start: 0, Add Exposure Days to Start: FALSE, Risk Window End: 30, Add Exposure Days to End: FALSE |

### Additional Population Settings

The final study population in which we will develop our model is a subset of the target cohort, because we may for example apply criteria that are dependent on the outcome, or we want to perform sensitivity analyses with sub-populations of the target cohort. For this we have to answer the following questions:

* ***What is the minimum amount of observation time we require before the start of the target cohort?***This choice could depend on the available patient time in the training data, but also on the time we expect to be available in the data sources we want to apply the model on in the future. The longer the minimum observation time, the more baseline history time is available for each person to use for feature extraction, but the fewer patients will qualify for analysis. Moreover, there could be clinical reasons to choose a short or longer look-back period.
* ***Can patients enter the target cohort multiple times?*** In the target cohort definition, a person may qualify for the cohort multiple times during different spans of time, for example if they had different episodes of a disease or separate periods of exposure to a medical product. The cohort definition does not necessarily apply a restriction to only let the patients enter once, but in the context of a particular patient-level prediction problem we may want to restrict the cohort to the first qualifying episode.
* ***Do we allow persons to enter the cohort if they experienced the outcome before?*** Do we allow persons to enter the target cohort if they experienced the outcome before qualifying for the target cohort? Depending on the particular patient-level prediction problem, there may be a desire to predict incident first occurrence of an outcome, in which case patients who have previously experienced the outcome are not at risk for having a first occurrence and therefore should be excluded from the target cohort. In other circumstances, there may be a desire to predict prevalent episodes, whereby patients with prior outcomes can be included in the analysis and the prior outcome itself can be a predictor of future outcomes.
* ***How do we define the period in which we will predict our outcome relative to the target cohort start?***We have to make two decisions to answer this question. First, does the time-at-risk window start at the date of the start of the target cohort or later? Arguments to make it start later could be that we want to avoid outcomes that were entered late in the record that actually occurred before the start of the target cohort or we want to leave a gap where interventions to prevent the outcome could theoretically be implemented. Second, we need to define the time-at-risk by setting the risk window end, as some specification of days offset relative to the target cohort start or end dates.
* ***Do we require a minimum amount of time-at-risk?*** We have to decide if we want to include patients that did not experience the outcome but did leave the database earlier than the end of our time-at-risk period. These patients may experience the outcome when we no longer observe them. For our prediction problem we decide to answer this question with “yes,” requiring a minimum time-at-risk for that reason. Furthermore, we have to decide if this constraint also applies to persons who experienced the outcome, or we will include all persons with the outcome irrespective of their total time at risk.

In our study, population settings are defined as described below:

***Population Settings #1***

|  |  |
| --- | --- |
| Item | Settings |
| minTimeAtRisk | 364 |
| requireTimeAtRisk | FALSE |
| addExposureDaysToStart | FALSE |
| riskWindowStart | 0 |
| washoutPeriod | 365 |
| addExposureDaysToEnd | FALSE |
| includeAllOutcomes | TRUE |
| priorOutcomeLookback | 99999 |
| binary | TRUE |
| removeSubjectsWithPriorOutcome | FALSE |
| riskWindowEnd | 30 |
| firstExposureOnly | FALSE |

## Statistical Analysis Method(s)

### Algorithms

In this study we will apply a Lasso Logistic Regression. Lasso logistic regression belongs to the family of generalized linear models, where a linear combination of the variables is learned and finally a logistic function maps the linear combination to a value between 0 and 1. The lasso regularization adds a cost based on model complexity to the objective function when training the model. This cost is the sum of the absolute values of the linear combination of the coefficients. The model automatically performs feature selection by minimizing this cost. We use the Cyclic coordinate descent for logistic, Poisson and survival analysis (Cyclops) package to perform large-scale regularized logistic regression: <https://github.com/OHDSI/Cyclops>.

### Model Evaluation

The following evaluations will be performed on the model:

|  |  |
| --- | --- |
| Evaluation | Description |
| Box Plots | The prediction distribution boxplots are box plots for the predicted risks of the people in the test set with the outcome (class 1: blue) and without the outcome (class 0: red). |
| Calibration Plot | The calibration plot shows how close the predicted risk is to the observed risk. The diagonal dashed line thus indicates a perfectly calibrated model. The ten (or fewer) dots represent the mean predicted values for each quantile plotted against the observed fraction of people in that quantile who had the outcome (observed fraction). The straight black line is the linear regression using these 10 plotted quantiles mean predicted vs observed fraction points. The two blue straight lines represented the 95% lower and upper confidence intervals of the slope of the fitted line. |
| Demographic Summary Plot | This plot shows for females and males the expected and observed risk in different age groups together with a confidence area. |
| Precision Recall Plot | The precision-recall curve is valuable for dataset with a high imbalance between the size of the positive and negative class. It shows the trade-off between precision and recall for different threshold. High precision relates to a low false positive rate, and high recall relates to a low false negative rate. High scores for both show that the classifier is returning accurate results (high precision), as well as returning a majority of all positive results (high recall). A high area under the curve represents both high recall and high precision. |
| Prediction Distribution Plots | The preference distribution plots are the preference score distributions corresponding to i) people in the test set with the outcome (red) and ii) people in the test set without the outcome (blue). |
| ROC Plot | The ROC plot plots the sensitivity against 1-specificity on the test set. The plot shows how well the model is able to discriminate between the people with the outcome and those without. The dashed diagonal line is the performance of a model that randomly assigns predictions. The higher the area under the ROC plot the better the discrimination of the model. |
| Smooth Calibration Plot | Similar to the traditional calibration shown above the Smooth Calibration plot shows the relationship between predicted and observed risk. the major difference is that the smooth fit allows for a more fine-grained examination of this. Whereas the traditional plot will be heavily influenced by the areas with the highest density of data the smooth plot will provide the same information for this region as well as a more accurate interpretation of areas with lower density. the plot also contains information on the distribution of the outcomes relative to predicted risk. However, the increased information game comes at a computational cost. It is recommended to use the traditional plot for examination and then to produce the smooth plot for final versions. |
| Test-Train Similarity Plot | The test-train similarity is presented by plotting the mean covariate values in the train set against those in the test set for people with and without the outcome. |
| Variable Scatter Plot | The variable scatter plot shows the mean covariate value for the people with the outcome against the mean covariate value for the people without the outcome. The size and colour of the dots correspond to the importance of the covariates in the trained model (size of beta) and its direction (sign of beta with green meaning positive and red meaning negative), respectively. |

## Quality Control

The PatientLevelPrediction package itself, as well as other OHDSI packages on which PatientLevelPrediction depends, use unit tests for validation. More information can be found in the Book of OHDSI at: <https://ohdsi.github.io/TheBookOfOhdsi/SoftwareValidity.html>

## Tools

To create the study package, ATLAS will be used to specify the cohorts, time-at-risk, covariate and population settings as well as which models will be analysed. Information on this is available in the Book of OHDSI at: <https://ohdsi.github.io/TheBookOfOhdsi/OhdsiAnalyticsTools.html#atlas>

The package developed in ATLAS will utilise the Patient-Level Prediction R package to run the analysis. More information on this is available at: <https://ohdsi.github.io/TheBookOfOhdsi/PatientLevelPrediction.html>

This study will be designed using OHDSI tools and run with R (20). More information about the tools can be found in the Appendix 'Study Generation Version Information'.

# Diagnostics

Reviewing the incidence rates of the outcomes in the target population prior to performing the analysis will allow us to assess its feasibility. The full table can be found in the 'Table and Figures' section under 'Incidence Rate of Target & Outcome'. Additionally, reviewing the characteristics of the cohorts provides insight into the cohorts being reviewed.

# Data Analysis Plan

## Algorithm Settings

***Model Settings Settings #1 - LassoLogisticRegressionSettings***

|  |  |
| --- | --- |
| Covariates | Settings |
| seed |  |
| variance | 0.01 |

## Covariate Settings

The covariates (constructed using records on or prior to the target cohort start date) are used within this prediction mode include the following. Each covariate needs to contain at least 0.001 subjects to be considered for the model.

***Covariate Settings #1***

|  |  |
| --- | --- |
| Covariates | Settings |
| VisitCountMediumTerm | FALSE |
| ObservationShortTerm | FALSE |
| shortTermStartDays | -30 |
| MeasurementRangeGroupShortTerm | FALSE |
| ConditionOccurrenceLongTerm | FALSE |
| DrugEraStartLongTerm | FALSE |
| VisitCountShortTerm | FALSE |
| Chads2Vasc | FALSE |
| ConditionGroupEraStartLongTerm | FALSE |
| ConditionEraShortTerm | FALSE |
| Dcsi | FALSE |
| DrugGroupEraLongTerm | TRUE |
| DrugGroupEraShortTerm | TRUE |
| ConditionEraStartLongTerm | FALSE |
| temporal | FALSE |
| DemographicsIndexMonth | FALSE |
| ConditionOccurrencePrimaryInpatientLongTerm | FALSE |
| ConditionEraAnyTimePrior | FALSE |
| addDescendantsToInclude | FALSE |
| ConditionGroupEraStartMediumTerm | FALSE |
| ProcedureOccurrenceLongTerm | FALSE |
| DrugExposureLongTerm | FALSE |
| DrugEraStartShortTerm | FALSE |
| DistinctIngredientCountMediumTerm | FALSE |
| DistinctMeasurementCountShortTerm | FALSE |
| MeasurementRangeGroupLongTerm | FALSE |
| ConditionGroupEraOverlapping | FALSE |
| MeasurementRangeGroupMediumTerm | FALSE |
| DrugGroupEraStartMediumTerm | FALSE |
| MeasurementAnyTimePrior | FALSE |
| MeasurementMediumTerm | FALSE |
| includedCovariateIds |  |
| ConditionOccurrenceAnyTimePrior | FALSE |
| DistinctConditionCountLongTerm | FALSE |
| MeasurementValueLongTerm | FALSE |
| DrugEraShortTerm | FALSE |
| DrugGroupEraAnyTimePrior | FALSE |
| DrugEraOverlapping | FALSE |
| ConditionOccurrencePrimaryInpatientAnyTimePrior | FALSE |
| ConditionEraMediumTerm | FALSE |
| ConditionEraOverlapping | FALSE |
| ConditionEraStartShortTerm | FALSE |
| ObservationAnyTimePrior | FALSE |
| VisitConceptCountShortTerm | FALSE |
| DemographicsEthnicity | FALSE |
| DistinctIngredientCountLongTerm | FALSE |
| ConditionOccurrencePrimaryInpatientShortTerm | FALSE |
| DemographicsAgeGroup | TRUE |
| DistinctProcedureCountShortTerm | FALSE |
| DistinctObservationCountMediumTerm | FALSE |
| includedCovariateConceptIds |  |
| DrugGroupEraStartShortTerm | FALSE |
| addDescendantsToExclude | FALSE |
| DrugEraLongTerm | FALSE |
| DistinctConditionCountShortTerm | FALSE |
| ConditionGroupEraShortTerm | TRUE |
| ConditionEraStartMediumTerm | FALSE |
| VisitCountLongTerm | FALSE |
| DemographicsRace | FALSE |
| ProcedureOccurrenceAnyTimePrior | FALSE |
| DistinctObservationCountLongTerm | FALSE |
| ProcedureOccurrenceMediumTerm | FALSE |
| CharlsonIndex | FALSE |
| DemographicsPriorObservationTime | FALSE |
| MeasurementShortTerm | FALSE |
| DistinctProcedureCountMediumTerm | FALSE |
| ConditionEraLongTerm | FALSE |
| DrugGroupEraStartLongTerm | FALSE |
| DemographicsGender | TRUE |
| DeviceExposureAnyTimePrior | FALSE |
| ObservationLongTerm | FALSE |
| DemographicsIndexYearMonth | FALSE |
| ConditionOccurrenceMediumTerm | FALSE |
| longTermStartDays | -365 |
| DemographicsAge | FALSE |
| DrugGroupEraOverlapping | FALSE |
| DistinctMeasurementCountLongTerm | FALSE |
| MeasurementRangeGroupAnyTimePrior | FALSE |
| DistinctConditionCountMediumTerm | FALSE |
| DrugGroupEraMediumTerm | FALSE |
| ProcedureOccurrenceShortTerm | FALSE |
| ObservationMediumTerm | FALSE |
| ConditionGroupEraAnyTimePrior | FALSE |
| Chads2 | FALSE |
| DrugExposureAnyTimePrior | FALSE |
| DeviceExposureLongTerm | FALSE |
| DemographicsTimeInCohort | FALSE |
| DistinctMeasurementCountMediumTerm | FALSE |
| MeasurementValueShortTerm | FALSE |
| DeviceExposureMediumTerm | FALSE |
| ConditionGroupEraStartShortTerm | FALSE |
| ConditionOccurrencePrimaryInpatientMediumTerm | FALSE |
| MeasurementLongTerm | FALSE |
| DemographicsIndexYear | FALSE |
| MeasurementValueMediumTerm | FALSE |
| DrugEraStartMediumTerm | FALSE |
| MeasurementValueAnyTimePrior | FALSE |
| DistinctObservationCountShortTerm | FALSE |
| DrugEraMediumTerm | FALSE |
| ConditionGroupEraLongTerm | TRUE |
| DrugExposureShortTerm | FALSE |
| DistinctIngredientCountShortTerm | FALSE |
| DeviceExposureShortTerm | FALSE |
| mediumTermStartDays | -180 |
| DemographicsPostObservationTime | FALSE |
| VisitConceptCountLongTerm | FALSE |
| VisitConceptCountMediumTerm | FALSE |
| excludedCovariateConceptIds |  |
| ConditionGroupEraMediumTerm | FALSE |
| DrugExposureMediumTerm | FALSE |
| DistinctProcedureCountLongTerm | FALSE |
| DrugEraAnyTimePrior | FALSE |
| endDays | -1 |
| ConditionOccurrenceShortTerm | FALSE |

***Covariate Settings #2***

|  |  |
| --- | --- |
| Covariates | Setting |
| VisitCountMediumTerm | FALSE |
| ObservationShortTerm | FALSE |
| shortTermStartDays | -30 |
| MeasurementRangeGroupShortTerm | FALSE |
| ConditionOccurrenceLongTerm | FALSE |
| DrugEraStartLongTerm | FALSE |
| VisitCountShortTerm | FALSE |
| Chads2Vasc | FALSE |
| ConditionGroupEraStartLongTerm | FALSE |
| ConditionEraShortTerm | FALSE |
| Dcsi | FALSE |
| DrugGroupEraLongTerm | FALSE |
| DrugGroupEraShortTerm | FALSE |
| ConditionEraStartLongTerm | FALSE |
| temporal | FALSE |
| DemographicsIndexMonth | FALSE |
| ConditionOccurrencePrimaryInpatientLongTerm | FALSE |
| ConditionEraAnyTimePrior | FALSE |
| addDescendantsToInclude | FALSE |
| ConditionGroupEraStartMediumTerm | FALSE |
| ProcedureOccurrenceLongTerm | FALSE |
| DrugExposureLongTerm | FALSE |
| DrugEraStartShortTerm | FALSE |
| DistinctIngredientCountMediumTerm | FALSE |
| DistinctMeasurementCountShortTerm | FALSE |
| MeasurementRangeGroupLongTerm | FALSE |
| ConditionGroupEraOverlapping | FALSE |
| MeasurementRangeGroupMediumTerm | FALSE |
| DrugGroupEraStartMediumTerm | FALSE |
| MeasurementAnyTimePrior | FALSE |
| MeasurementMediumTerm | FALSE |
| includedCovariateIds |  |
| ConditionOccurrenceAnyTimePrior | FALSE |
| DistinctConditionCountLongTerm | FALSE |
| MeasurementValueLongTerm | FALSE |
| DrugEraShortTerm | FALSE |
| DrugGroupEraAnyTimePrior | FALSE |
| DrugEraOverlapping | FALSE |
| ConditionOccurrencePrimaryInpatientAnyTimePrior | FALSE |
| ConditionEraMediumTerm | FALSE |
| ConditionEraOverlapping | FALSE |
| ConditionEraStartShortTerm | FALSE |
| ObservationAnyTimePrior | FALSE |
| VisitConceptCountShortTerm | FALSE |
| DemographicsEthnicity | FALSE |
| DistinctIngredientCountLongTerm | FALSE |
| ConditionOccurrencePrimaryInpatientShortTerm | FALSE |
| DemographicsAgeGroup | TRUE |
| DistinctProcedureCountShortTerm | FALSE |
| DistinctObservationCountMediumTerm | FALSE |
| includedCovariateConceptIds |  |
| DrugGroupEraStartShortTerm | FALSE |
| addDescendantsToExclude | FALSE |
| DrugEraLongTerm | FALSE |
| DistinctConditionCountShortTerm | FALSE |
| ConditionGroupEraShortTerm | FALSE |
| ConditionEraStartMediumTerm | FALSE |
| VisitCountLongTerm | FALSE |
| DemographicsRace | FALSE |
| ProcedureOccurrenceAnyTimePrior | FALSE |
| DistinctObservationCountLongTerm | FALSE |
| ProcedureOccurrenceMediumTerm | FALSE |
| CharlsonIndex | FALSE |
| DemographicsPriorObservationTime | FALSE |
| MeasurementShortTerm | FALSE |
| DistinctProcedureCountMediumTerm | FALSE |
| ConditionEraLongTerm | FALSE |
| DrugGroupEraStartLongTerm | FALSE |
| DemographicsGender | TRUE |
| DeviceExposureAnyTimePrior | FALSE |
| ObservationLongTerm | FALSE |
| DemographicsIndexYearMonth | FALSE |
| ConditionOccurrenceMediumTerm | FALSE |
| longTermStartDays | -365 |
| DemographicsAge | FALSE |
| DrugGroupEraOverlapping | FALSE |
| DistinctMeasurementCountLongTerm | FALSE |
| MeasurementRangeGroupAnyTimePrior | FALSE |
| DistinctConditionCountMediumTerm | FALSE |
| DrugGroupEraMediumTerm | FALSE |
| ProcedureOccurrenceShortTerm | FALSE |
| ObservationMediumTerm | FALSE |
| ConditionGroupEraAnyTimePrior | FALSE |
| Chads2 | FALSE |
| DrugExposureAnyTimePrior | FALSE |
| DeviceExposureLongTerm | FALSE |
| DemographicsTimeInCohort | FALSE |
| DistinctMeasurementCountMediumTerm | FALSE |
| MeasurementValueShortTerm | FALSE |
| DeviceExposureMediumTerm | FALSE |
| ConditionGroupEraStartShortTerm | FALSE |
| ConditionOccurrencePrimaryInpatientMediumTerm | FALSE |
| MeasurementLongTerm | FALSE |
| DemographicsIndexYear | FALSE |
| MeasurementValueMediumTerm | FALSE |
| DrugEraStartMediumTerm | FALSE |
| MeasurementValueAnyTimePrior | FALSE |
| DistinctObservationCountShortTerm | FALSE |
| DrugEraMediumTerm | FALSE |
| ConditionGroupEraLongTerm | FALSE |
| DrugExposureShortTerm | FALSE |
| DistinctIngredientCountShortTerm | FALSE |
| DeviceExposureShortTerm | FALSE |
| mediumTermStartDays | -180 |
| DemographicsPostObservationTime | FALSE |
| VisitConceptCountLongTerm | FALSE |
| VisitConceptCountMediumTerm | FALSE |
| excludedCovariateConceptIds |  |
| ConditionGroupEraMediumTerm | FALSE |
| DrugExposureMediumTerm | FALSE |
| DistinctProcedureCountLongTerm | FALSE |
| DrugEraAnyTimePrior | FALSE |
| endDays | 0 |
| ConditionOccurrenceShortTerm | FALSE |

## Model Development & Evaluation

To build and internally validate the models, we will partition the labelled data into a train set (75%) and a test set (25%).

The hyper-parameters for the models will be assessed using 3-fold cross validation on the train set and a final model will be trained using the full train set and optimal hyper-parameters.

The internal validity of the models will be assessed on the test set. The external validity of the models with be assessed on recent COVID-19 data. We will use the area under the receiver operating characteristic curve (AUC) to evaluate the discriminative performance of the models and plot the predicted risk against the observed fraction to visualize the calibration. See 'Model Evaluation' section for more detailed information about additional model evaluation metrics.

## Analysis Execution Settings

There are is 1 target cohorts evaluated for 2 outcomes over 1 model over 2 covariates settings and over 1 population setting. In total there are 4 analysis performed.

# Strengths & Limitations

Strength

* The analysis can help gain insight into the clinical usefulness of each developed model by identifying whether it is transportable.

Limitations

* The external validation datasets may not have a sufficient number of outcomes to be used in the analysis.
* Although the CDM standardizes the vocabularies of the datasets, the concept recording distributions are likely to differ between databases and it is unknown how much this will limit model transportability.

# Protection of Human Subjects

For this study, participants from various countries will process personal data from individuals which is collected in national/regional electronic health record databases. Due to the sensitive nature of this personal medical data, it is important to be fully aware of ethical and regulatory aspects and to strive to take all reasonable measures to ensure compliance with ethical and regulatory issues on privacy.

In agreement with these regulations, rather than combining person level data and performing only a central analysis, local analyses will be run, which generate non-identifiable aggregate summary results.

All the databases used in this study have a well-developed mechanism to ensure that regulations dealing with ethical use of the data and adequate privacy control are adhered to.

If required, the protocol has been reviewed by the Institutional Review Boards of the respective databases.

# Plans for Disseminating & Communicating Study Results

Dissemination activities to be undertaken will have mainly, although not exclusively, a scientific nature (articles, presentations at conferences, etc.).

# Tables & Figures

## Incidence Rate of Target & Outcome

***Feasibility assessment training data***

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| T | O | TAR start | TAR end | T size | O count | O incidence |

***Feasibility assessment validation data***

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Database | O | TAR | T size | O count | O incidence |

# Appendices

## Study Generation Version Information

Skeleton Version: PatientLevelPredictionStudy - v0.0.1

Identifier / Organization: OHDSI

# References

1. World Health Organization. WHO Director-General's opening remarks at the media briefing on COVID-19. 2020 March 11 [cited 2020 March 28]. Available from: <https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020>.
2. Wang CJ, Ng CY, Brook RH. Response to COVID-19 in Taiwan: Big Data Analytics, New Technology, and Proactive Testing. JAMA. 2020.
3. Yang X, Yu Y, Xu J, Shu H, Liu H, Wu Y, et al. Clinical course and outcomes of critically ill patients with SARS-CoV-2 pneumonia in Wuhan, China: a single-centered, retrospective, observational study. The Lancet Respiratory Medicine. 2020.
4. World Health Organization. Clinical management of severe acute respiratory infection (‎SARI)‎ when COVID-19 disease is suspected: interim guidance, 13 March 2020. Geneva: World Health Organization; 2020.
5. Emanuel EJ, Persad G, Upshur R, Thome B, Parker M, Glickman A, Zhang C, Boyle C, Smith M, Phillips JP. Fair allocation of scarce medical resources in the time of Covid-19.
6. Grasselli G, Pesenti A, Cecconi M. Critical Care Utilization for the COVID-19 Outbreak in Lombardy, Italy: Early Experience and Forecast During an Emergency Response. JAMA. 2020 Mar 13.
7. Fine MJ, Auble TE, Yealy DM, Hanusa BH, Weissfeld LA, Singer DE, Coley CM, Marrie TJ, Kapoor WN. A prediction rule to identify low-risk patients with community-acquired pneumonia. New England journal of medicine. 1997 Jan 23;336(4):243-50.
8. Kohno S, Seki M, Takehara K, Yamada Y, Kubo K, Ishizaka A, Soma K. Prediction of requirement for mechanical ventilation in community-acquired pneumonia with acute respiratory failure: a multicenter prospective study. Respiration. 2013;85(1):27-35.
9. Jones BE, Jones J, Bewick T, Lim WS, Aronsky D, Brown SM, Boersma WG, van der Eerden MM, Dean NC. CURB-65 pneumonia severity assessment adapted for electronic decision support. Chest. 2011 Jul 1;140(1):156-63.
10. Tashiro M, Fushimi K, Takazono T, Kurihara S, Miyazaki T, Tsukamoto M, Yanagihara K, Mukae H, Tashiro T, Kohno S, Izumikawa K. A mortality prediction rule for non-elderly patients with community-acquired pneumonia. BMC pulmonary medicine. 2016 Dec;16(1):39.
11. Uematsu H, Kunisawa S, Sasaki N, Ikai H, Imanaka Y. Development of a risk-adjusted in-hospital mortality prediction model for community-acquired pneumonia: a retrospective analysis using a Japanese administrative database. BMC pulmonary medicine. 2014 Dec;14(1):203.
12. Ahn JH, Choi EY. Expanded A-DROP score: a new scoring system for the prediction of mortality in hospitalized patients with community-acquired pneumonia. Scientific reports. 2018 Oct 1;8(1):1-9.
13. Charles PG, Wolfe R, Whitby M, Fine MJ, Fuller AJ, Stirling R, Wright AA, Ramirez JA, Christiansen KJ, Waterer GW, Pierce RJ. SMART-COP: a tool for predicting the need for intensive respiratory or vasopressor support in community-acquired pneumonia. Clinical infectious diseases. 2008 Aug 1;47(3):375-84.
14. Guo L, Wei D, WU Y, ZHOU M, ZHANG X, Li Q, Qu J. Clinical features predicting mortality risk in patients with viral pneumonia: the MuLBSTA score. Frontiers in Microbiology. 2019;10:2752.
15. Gong J, Ou J, Qiu X, Jie Y, Chen Y, Yuan L, Cao J, Tan M, Xu W, Zheng F, Shi Y. Multicenter Development and Validation of a Novel Risk Nomogram for Early Prediction of Severe 2019-Novel Coronavirus Pneumonia. Available at SSRN 3551365. 2020 Mar 9.
16. Lu J, Hu S, Fan R, Liu Z, Yin X, Wang Q, Lv Q, Cai Z, Li H, Hu Y, Han Y. ACP risk grade: a simple mortality index for patients with confirmed or suspected severe acute respiratory syndrome coronavirus 2 disease (COVID-19) during the early stage of outbreak in Wuhan, China.
17. Reps JM, Schuemie MJ, Suchard MA, Ryan PB, Rijnbeek PR. Design and implementation of a standardized framework to generate and evaluate patient-level prediction models using observational healthcare data. Journal of the American Medical Informatics Association. 2018 Aug;25(8):969-75.
18. Steyerberg EW, Moons KGM, van der Windt DA, Hayden JA, Perel P, Schroter S, et al. Prognosis Research Strategy (PROGRESS) 3: prognostic model research. PLoS Med. 2013;10(2):e1001381.
19. Moons KGM, Altman DG, Reitsma JB, Ioannidis JPA, Macaskill P, Steyerberg EW, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. Ann Intern Med. 2015;162(1):W1-73.
20. Team RC. R: A language and environment for statistical computing. 2013.