

Early detection of COVID-19 mortality risk using non-invasive clinical characteristics

Mahdi Mahdavi

School of Medicine, Shahid Beheshti University of Medical Sciences,

Hadi Choubdar

School of Medicine, Shahid Beheshti University of Medical Sciences,

Erfan Zabeh

Department of Biomedical Engineering, Columbia University

Michael Rieder

Department of Medicine, Schulich School of Medicine and Dentistry, University of Western Ontario

Safieddin Safavi-Naeini

Centre for Intelligent Antenna and Radio Systems at the university (CIARS)

Vida Khanlarzadeh

University of Western Ontario

Zsolt Jobbagy

Department of Pathology, Immunology and Molecular Pathology, Rutgers New Jersey Medical School

Amirata Ghorbani

Department of Electrical Engineering, Stanford University <https://orcid.org/0000-0001-6465-2321>

Atefeh Abedini

Chronic Respiratory Diseases Research Center, National Research Institute of Tuberculosis and Lung Diseases (NRITLD), Shahid Beheshti University of Medical Sciences,

Arda Kiani

Tracheal Diseases Research Center, National Research Institute of Tuberculosis and Lung Diseases (NRITLD), Shahid Beheshti University of Medical Sciences and Health Services

Reza Lashgari

Brain Engineering Research Center, Institute for Research in Fundamental Sciences (IPM)

Ehsan Kamrani (✉ ekamrani@gmail.com)

Dept. of Electrical & Computer Engineering, University of Waterloo <https://orcid.org/0000-0002-0919-1179>

Article

Keywords: COVID-19, mortality risk prediction, non-invasive biomarkers, models

Posted Date: October 7th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-86363/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

With no effective treatment currently available and maximum preventive measures already in place, more interventions in the clinical field are needed to decrease COVID-19 patient mortality. Early prediction of mortality risk in COVID-19 patients can decrease mortality by assuring efficient resource allocation and treatment planning. This study conducts an early prediction of COVID-19 prognosis using laboratory, clinical, and demographic data collected from patients in the first day of admission. Three machine learning models were developed to investigate and compare the prediction power of invasive and noninvasive biomarkers. The results suggest that early mortality prediction of patients via non-invasive biomarkers provides significant accuracy and can be used as a triage assisting tool without the need for additional costs or waiting time of laboratory tests.

Introduction

The SARS-COV-2 pandemic has tremendously strained economic and healthcare infrastructures worldwide, leaving a trail of more than 850,000 deaths behind as of September 2, 2020 [1]. With no effective treatment or vaccine, an average global death rate of around 6000 per day could lead to the death of approximately

2.2 million individuals in one year. Furthermore, the upcoming concomitance of seasonal Influenza and COVID-19 [2] will only worsen the situation. Even though strict social distancing and preventive measures are still in effect, the global mortality and prevalence curve of the disease shows little improvement [1]. As a result, more focus on the clinical aspect of the disease is needed to reduce patient mortality rates. A crucial clinical aspect of the COVID-19 outbreak is the possible outcomes of the disease. The severity spectrum of COVID-19 varies from mild asymptomatic to severe and critical conditions [3]. Critical patients will need timely intensive care unit (ICU) admission and ventilators. In China, it has been reported about 54% of critical patients were unable to receive timely ICU care, and 30% of patients who died did not receive mechanical ventilation in time [4]. With large patient loads, exhausted medical personnel, and insufficient medical resources, expedited identification of patients that have high mortality risks becomes a key factor in decreasing patient deaths.

Physicians are often unable to accurately predict the prognosis of COVID-19 patients upon their admission until later stages of the disease. Furthermore, the course of COVID-19 can take unpredictable turns where the condition of a seemingly stable patient deteriorates rapidly to a critical state [5]; this could catch even the most skilled physicians off guard. To enhance clinical prediction, Artificial Intelligence (AI) models could be valuable assistants since they can detect complex patterns in large datasets [6] [7].

The main question of this study was framed as a classification problem. The binary outcome was expressed as patient expiration or discharge. Patient demographic, clinical, and laboratory data were used as inputs. Several machine learning prognostic models have been proposed for Covid19 mortality

[8] [9] [10] [11]. However, there is no comprehensive study that evaluated and compared the prediction power of non-invasive and invasive. In this study, predictive features were divided into invasive laboratory tests and non-invasive demographic and clinical features. Three machine learning models were developed to investigate and compare the prediction power of these feature groups; two using each of these groups and one using both (Figure 1). The first 48 hours of patients' admission is crucial for various clinical decisions (e.g., whether to request additional care equipment for the patient). Furthermore, it has been reported that many COVID-19 patients experienced their first exacerbation period 24 to 48 hours after admission [12]. Accordingly, we based our model on data from the first day of patients' admission to provide a tool that can be beneficial in real-life environment.

The joint model displayed satisfactory performance; this model could provide rapid, accurate mortality prediction using features that are routinely collected during the first day of patient admission. Furthermore, our non-invasive model displayed good prediction accuracy and could be utilized as a triage tool to deliver an unbiased initial mortality risk estimation. Dimension reduction analysis using the Neighborhood Component Analysis (NCA) method revealed that age, oxygen saturation (SPO₂), partial thromboplastin time (PTT), lactate dehydrogenase (LDH), and Aspartate transaminase (AST) were among the most contributing features for mortality prediction. To investigate the temporal performance of classifiers, we compared model prediction accuracies across patient expiration intervals (i.e., number of days between admission and expiration). This analysis across the expiration time spectrum revealed that although the invasive model displays a proper performance for imminent deaths, it was outperformed by non-invasive model over longer expiration intervals.

Results

Data resources

Electronic medical records of 328 patients who were admitted to Masih Daneshvari Hospital between February 20th, 2020, and May 4th, 2020, were initially included. Patient diagnosis and severity classification was carried out using the criteria presented in Table 1. After exclusion of 85 patients, data from 243 individuals (62.6% male, 37.4% female) were used for model development (Supplementary Figure S1). The median age of the study population was 62 (22). Furthermore, 184 (75.7%) patients were documented as severe, and 59 (24.3%) as non-severe cases. Dyspnea (88.1%), cough (83.4%), and fever (66.9%) were the three most frequent symptoms. Hypertension (42.4%), diabetes mellitus (33.7%) and cardiovascular disease (15.2%) were the three most frequent comorbidities among patients (Table 2).

Contribution of individual factors to mortality prediction

Preliminary analysis of single parameters and earlier reports of COVID-19 mortality prediction [8][13], suggested that mortality outcome can be explained by the value of individual biomarkers. However, outcome prediction through single biomarker is not highly accurate. The key idea for increasing the prediction accuracy is to combine the information of individual features together. One common technique

in medical data analysis is generation of decision trees. However, this study utilized binary classifiers which can provide more interpretable results. Accordingly, to use the information of all biomarkers together, the classifying problem was transferred from one dimensional space (categorizing based on a single variable) into a high dimensional space where the space axes are built by independent predictors. Panel B of Figure 2 illustrates a proper two-dimensional space where the discharged and expired cases can be discriminated. In contrast, panel C illustrates a non-discriminative feature subspace. To compare the contribution of individual features to the prediction of mortality outcomes, Neighbourhood Component Analysis (NCA) was implemented; this non-parametric analysis is an embedded dimensionality reduction method for selecting features with the goal of maximizing prediction accuracy of classification algorithms. Figure 2 Panel A displays the weights of fitted NCA model for classification of mortality outcome using 37 pruned biomarkers. Features with weights close to zero have less contribution in prediction of the final outcome. The results of dimensional reduction analysis suggest that several predictors, including SPO_2 , age, PTT, LDH, and AST, have significant weights in mortality prediction. As is shown in Figure 2, these significant predictors include both invasive (gray shaded) and non-invasive (green shaded) biomarkers. The presence of non-invasive biomarkers as significant mortality predictors motivated us to investigate the possibility of predicting mortality risk using only non-invasive biomarkers.

Predictive capability of invasive and non-invasive model

Following the results of dimensionality reduction analysis, we investigated the possibility of accurate mortality prediction via non-invasive and invasive biomarkers and their combination. Towards this aim, mortality prediction was implemented using three models (Figure 1). The joint model utilized all the demographic, laboratory, and clinical biomarkers as inputs; this model is considered ideal, where all the required biomarkers are present for prediction. To investigate the differences between invasive and non-invasive features for outcome prediction, two separate models were developed; one solely based on laboratory features (shown by grey color in Figures) and the other only based on non-invasive features (shown by green color in Figures). For each of the Joint, invasive, and non-invasive models, a linear support vector machine (SVM) algorithm was trained. To generalize the results of model prediction to an independent data set, we performed 10-fold cross validation on preprocessed patient data. For each model, the accuracy reported on the test data-set and the reported classification loss obtained by the cross-validated kernel classification model [14].

To better validate the performance of the predictive algorithms, independent of the algorithm decision criteria, receiver operating characteristic (ROC) curves were generated. The ROC curve is a graphical plot that illustrates the predictive ability of a binary classifier system as its discrimination threshold is varied. Figure 3 suggest that not only the prediction power of non-invasive model is comparable with the joint model, but also outperforms the invasive model. Thus, the non-invasive model can be confidently used as an assistive triage tool to sieve patients with high mortality risks.

Earlier studies of COVID-19 mortality predictions mostly focused on invasive biomarkers [8][11]. Evaluating invasive biomarkers provides more direct and causal inferences about our physiological state. In contrast, non-invasive features contain broader, indirect information about the body. Thus, it could be hypothesized that accurate mortality risk anticipation is plausible with a sufficient number of non-invasive features.

Thus, we aimed to investigate whether the predictive power of non-invasive features (population signal) are comparable to those of invasive features (causal signals) independent of the fine-tuned combinations of measurements specific to this study. Although the ROC analysis tackles this question (suggesting the superiority of non-invasive models) it is still unclear whether the excellence of non-invasive model is due to the higher absolute information that individual non-invasive measurements carry, comparing to invasive biomarkers or the specific combination of the current measurements.

To determine which of the invasive or non-invasive features are more informative, the number of predictors (i.e., model dimension) for the SVM model were fixed. Then, new invasive and non-invasive models were generated with corresponding fixed model dimensions.(Figure 3, panel B). This procedure was repeated over 40 trials.The average model accuracy over trials (AUC score) is illustrated in Figure 3 for each group of n-predictor models (invasive model: n = 1:28, non-invasive model: n = 1:11). The shaded error areas in 3 represent Standard Error of Mean (SEM). As expected, the prediction variance was higher in low-dimension models. The comparison between gray and green curves verified that non-invasive features are more informative compared to laboratory biomarkers (Wilcox test P-value<0.01).

Comparison of the prediction horizon

According to the functional dissimilarities that we observed between invasive and non-invasive models, we hypothesized whether the temporal range of prediction differed between models. To evaluate the prediction range of each model, the data was splitted into 12 distinct time intervals based on the expiration date of patients. The distribution of patient expiration dates in Figure 4, panel B, indicates that most of the patients expired within the first week of admission (median expiration interval = 6 days). Training independent SVM classifiers for each of the time intervals gives statistically invalid results as the patient sample size is already reduced after splitting procedure. To resolve this issue, we kept previous invasive and non-invasive models which each of them were trained over the complete dataset and then we calculated the fraction of false death prediction to the total number of predictions within each of the days to outcome intervals (F/TF). In Figure 4 panel A, we define the prediction performance as $1 - F/TF$, and then we compare it between invasive and non-invasive models.

The comparison between green (non-invasive) and gray (invasive) traces in panel A. suggests that invasive features are more predictive within the last days prior to expiration. Moreover, the performance of the invasive predictor model declined as the interval between admission and expiration increased. It should be noted that the performance of non-invasive model was better over longer expiration intervals. The stability in prediction performance of non-invasive model suggests that demographic and clinical features can predict mortality risk independent of the date of their measurement.

Discussion

Conducted for the first time on Iranian patients, this study provides three practical prognostic models using invasive and non-invasive data from the first day of patients' admission to predict the COVID-19 mortality. Furthermore, the prediction power of non-invasive and invasive feature groups was evaluated across the temporal and feature number spectrum to reveal interesting result. Although invasive features are good predictors for the imminent future, they are outperformed by non-invasive features for a more distant future. Moreover, compared with the invasive model, the non-invasive model could provide better performance in higher, equal feature dimensions.

Predicting the trajectory destination of COVID-19 could provide substantial support for decreasing mortality rates. Rapid disease transmission and high patient load could quickly overload healthcare infrastructures; an overloaded medical system can result in higher mortality rates due to inefficient management of limited medical resources and personnel; this issue was highlighted by a study indicating that 30% of Chinese COVID-19 patients died without receiving ventilator support [4]. Furthermore, strict preventive measures, social isolation, and the distress caused by the diagnosis of the disease could lead to the activation of psychological defensive behaviors in patients where they underestimate their symptoms and do not seek immediate medical assistance [15]. This optimistic bias could be fatal if the condition of a patient suddenly worsens towards a critical stage. Similar to an early warning system, our models could amend these issues by providing unbiased, rapid prognosis prediction to support proper resource allocation and decision making.

Time is an important element in the fight against COVID-19. The disease has an unpredictable trajectory where the condition of some patients suddenly becomes critical [5], surprising even the most skilled physicians; this hampers physicians' performance by limiting their action time window. Furthermore, it has been shown that patients who later become critically ill carry significantly more viral loads even before their condition becomes critical [16]. Thus, rapid isolation of high-risk patients is required to decrease infection spread. Our models could alleviate these problems; by providing a prognosis prediction after the first day of patient's admission, the time window of clinical actions (e.g., requesting additional ventilators, registering patients on clinical trials) could be significantly increased, potentially decreasing mortality rates by enabling a more thoughtful approach by physicians.

We developed three predictive models using invasive features, non-invasive features, and both. Our joint model provides rapid, accurate predictions using features that are routinely collected upon patient admission, making it implementable even in conditions where imaging or sophisticated laboratory equipment is unavailable. Our results revealed that non-invasive features displayed good prediction capacity compared with the joint model (Figure 3, panel A). Furthermore, invasive features displayed more prediction accuracy for imminent deaths. In contrast, non-invasive features provided more prediction power to predict deaths that were further from admission day (Figure 4, panel A). This difference in prediction range might stem from the fact that invasive and non-invasive biomarkers have distinct temporal dynamicity and information content. Many key laboratory features, such as LDH and

PTT, have high temporal dynamicity; to maintain homeostasis, after an insulting event, these features tend to rise and then return to their normal range in a relatively short time [17]. Furthermore, laboratory feature abnormalities only show disruptions in body systems, which they are linked to, limiting their information content. The aforementioned factors limit the predictive temporal range of models, which are mainly based on laboratory biomarkers. Many non-invasive features, such as age or presence of comorbidity, can be seen as signals that contain a significant amount of compressed, less variable data. Physicians, by instinct and training, could decompress these signals to a certain degree. For example, the presence of diabetes is informative of persistent glucose metabolism, immune response, and vascular function distortions [18]; these distortions, to varying degrees, are continuously present even in patients with proper disease control.

Although laboratory features provide valuable prognostic information, their analysis requires invasive sampling. Many patients are wary of blood sampling [19]. Moreover, high patient load and equipment shortage could hinder the availability and accuracy of blood testing [20]; many biomarkers, such as LDH and blood gas tests, require careful sampling, preservation, and transportation to avoid errors resulting from complications, such as sample hemolysis or clotting [21]. Besides, many smaller health centers do not have access to laboratory equipment. Lab tests are also generally expensive. A study from the United States indicates that, even in the absence of a pandemic state, over 20% of patient medical care was not needed [22]. These rates would most likely increase considerably in a pandemic state; physicians, faced with a novel disease and no coherent guidelines, will request more unnecessary blood tests, patient admissions, and referrals. These unnecessary cares will impose a significant financial burden on patients and healthcare systems. Rapid triage of patients is also a critical factor, required to manage high patient loads [23]. However, an important downside of routine rapid triage in a pandemic situation is the increased mortality rate due to missing high-risk patients [11]. These patients might incorrectly be identified as mild and, without further workup, be advised to take a home-treatment approach. Our model using non-invasive features could provide fast, accurate prognosis prediction to augment the initial triage and avoid missing high-risk patients. To facilitate this triage approach, vital signs could be easily measured by wearable medical devices [24], and history data could be easily asked from patients by predefined questions.

Five prominent features were highlighted by NCA analysis in this study; age, SPO₂, LDH, AST, and PTT. Previous studies have shown that older age is positively associated with increased mortality in hospitalized COVID-19 patients [25]. Older age is associated with more infection susceptibility and an atypical response to viral pathogens due to reduced expression of type I interferon-beta [26]. Furthermore, age-related impairment of lymphocyte function along with abnormal expression of type 2 cytokines leads to prolonged pro-inflammatory responses; this weakens the host response to viral replication causing poor clinical outcomes and higher mortality [27]. In contrast to typical types of pneumonia, the initial phases of COVID-19 have little apparent symptoms, such as dyspnea. The cause is the fact that there is still carbon dioxide exchange through alveoli at these stages. However, the oxygen exchange is disturbed due to the alveolar collapse. This type of hypoxia, called “silent hypoxia,” leads to the progression of pneumonia in the absence of clinical symptoms [28]. It also causes a vicious cycle, where hypoxia

promotes the activity of the local inflammatory system causing further damage and higher hypoxia [29]. therefore, SPO2 could be a decisive factor to uncover the pneumonia progression and the severity state of patients.

Pulse oximetry, via wearable devices or hospital equipment, could show decreased levels of SPO2, in blood; this is valuable for early detection of the hypoxemia. Elevated levels of LDH could reflect tissue injury caused by SARS-CoV-2 and concurrent lung fibrosis. Indeed, abnormal LDH is commonly seen in idiopathic lung fibrosis [8][30]. Furthermore, a robust immune response to SARS-CoV-2 infection and subsequent cytokine storm could cause multi-organ damage, which causes a further rise in the LDH level [31]. Another organ which is affected by the cytokine storm is the liver, even though, it is not a primary target for SARS-COV-2. As a result, abnormal levels of liver function biomarkers such as AST could be a manifestation of severe disease and poor outcomes [31]. The inflammatory response promoted by severe SARS-CoV-2 infection could cause endothelial damage, distortion of the coagulation cascade function, and coagulopathy. Therefore, levels of PTT, a coagulation biomarker, during COVID-19 infection can be informative of coagulopathy progression and disease severity [32].

Limitations

The results of this study should be interpreted in light of several limitations. This study was carried out within a retrospective framework. Consequently, supervision was not possible to increase the quality of data documentation when patients were admitted. Furthermore, the analysis interval of this study encompassed the first disease wave. Thus, medical records were documented in haste as high patient loads and limited medical staff forced the medical system to prioritize patient treatment. Researchers were not blind to outcomes. No external validation data was utilized due to limits imposed by the pandemic state of hospitals and preventive regimes. The Massih Daneshvari Hospital had more severe and expired patients since it was a primary care center for COVID-19. Finally, qualitative CRP, a feature reported by several studies to be associated with disease severity, was removed from the analysis due to high missing values due to limited laboratory resources and incomplete medical records caused by the pandemic. However, with the presence of other acute phase reactants and inflammatory markers, such as ESR, platelet number, and LDH, in this model, it is likely that a significant portion of variance explained by CRP was compensated by these features.

Future works

To increase speed and convenience, imaging features were not utilized in this study. Future works can compare the prediction power of imaging features with laboratory and non-invasive features. This study, conducted as a pilot study, was not externally validated. Future studies could include data from other hospitals for external validation. Future projects can expand the practicality of our study by devising prognosis prediction software on various platforms. In this study, binary outcome (i.e., discharged and expired) was used as outcome. Future projects can focus on other outcomes, such as whether a patient was intubated or admitted to ICU as outcomes. To devise specific prognostic models, Future studies can focus on individual groups of comorbidities (e.g., cardiovascular) to develop separate models. Finally,

continuous data input from various hospitals could be used to develop and incrementally train an online learning model to predict the prognosis of COVID-19 patients, giving increasingly precise and updated results to be used in clinical and non-clinical settings.

Conclusion

Prediction of mortality outcomes during the COVID-19 pandemic is an important concept which can reduce disease mortality rates by giving us insights into where and with whom to intervene. Prediction capacity of laboratory biomarkers are distinct from those of clinical and demographic data. To investigate these differences in this study, predictor features were divided into invasive laboratory tests and non-invasive demographic and clinical data.

Three prognostic machine learning models were developed using routine invasive and non-invasive biomarkers from the first day of patients' admission; two using each of these groups and one using both. The models displayed optimal prediction performance, making them valuable assistive tools in different settings for clinical decision making and resource allocation. Furthermore, the implemented non-invasive model can be used for rapid, accurate triage of patients without the need for additional costs or waiting time of laboratory tests.

Analysis of invasive and non-invasive models across equal feature numbers revealed that non-invasive model, irrespective of feature choice, provided better performance and could surpass the invasive model at a certain number of features. This adds weight to the hypothesis that, given enough features, the information content of non-invasive factors could provide mortality prediction which are as good as laboratory models. Temporal analysis across patient expiration intervals revealed that while invasive biomarkers are better predictors for near future compared with non-invasive features, they were outperformed by non-invasive biomarkers over larger expiration temporal intervals.

Methods

To increase the reporting quality and clarification, we aimed to follow the Transparent Reporting of a multi-variable prediction model for individual prognosis or diagnosis (TRIPOD) reporting style to the extent that pandemic situational limitations permitted.

Ethics

The protocol for this study was approved by the Ethics Committee of Shahid Beheshti University of Medical sciences. Patient data were anonymized before the data analysis phase.

Study setting and population

This retrospective study was carried out using archived electronic medical records of COVID-19 patients who were admitted to the Masih Daneshvari Hospital in Tehran, Iran. As Iran's largest respiratory and

pulmonary care center, the Masih Daneshvari Hospital was one of the first medical centers that admitted COVID-19 patients early during the pandemic.

The inclusion criterion for this study was hospital admission due to the initial diagnosis of COVID-19 infection by a physician according to the 5th Iranian COVID-19 guideline (Table 1). To enter this study, 328 patients who were admitted between February 20th, 2020, to May 4th, 2020, were randomly selected from the hospital's patient data pool. The review of medical records was initiated on June 5th. Medical records were reviewed separately by two physicians. Two pulmonologists adjudicated discordances. Afterwards, 8 patients who left the hospital against medical advice with consent, 22 patients with uncertain or rolled out COVID-19 diagnosis, 25 patients who were referred from other hospitals, 3 patients with age under 18, 14 patients with more than 20% missing data, 12 patients received radically different treatment protocols (i.e., were enrolled in clinical trials) and 1 patient who had a cardiac arrest shortly after arrival to the emergency ward, were excluded. No patients with breastfeeding or pregnancy were present in the initially selected population. As a result, 115 patients with an "expired" outcome (47.3%) and 128 patients with "discharged" outcome (52.7) were included for model development (Supplementary Figure S1).

Definition of variables

Data from the first 24 hours of patients' admission was used in this study. The initial data, comprised of 57 features, was categorized into two groups of feature; demographic and patient history features were labeled as non-invasive group, and laboratory results were labeled as invasive group. Demographic and history features were extracted from admission history, medical progress notes, and nursing notes. Laboratory features were extracted from the results of blood tests, which were ordered by physicians during the first 24 hours of admission.

The severity of a patient's condition was evaluated by physicians upon admission according to the criteria presented in Table 1. With the large magnitude of the outbreak and limited ICU beds, the severity criteria were tuned to the hospital's patient load and equipment to provide an efficient containment response. All patient vital signs were measured via electronic patient monitoring devices by a physician or nurse. Blood samples were analyzed in the central medical laboratory of Masih Daneshvari Hospital.

To predict mortality outcomes, the framework of this study was expressed as a classification problem. Two outcome classes were defined; the discharged group (outcome = 1) consisted of COVID-19 patients that were discharged after the completion of their treatment and two consecutive negative PCR results. The expired group (outcome = 0) consisted of patients who died at any point during their treatment course.

Statistical Tests

The Kolmogorov–Smirnov test was used to examine distribution normality. Mean and standard deviation were used to describe normal continuous variables. Median and interquartile range were used to describe

non-normal Continuous variables. Categorical data were expressed as frequency in percent. Kruskal-Wallis and Fisher's exact test tests were used to test significance for continuous and categorical variables, respectively. standardization (Z-score normalization) was used for feature scaling. Initial statistical analysis was performed using SPSS 26.0 (IBM Corp. Released 2019. IBM SPSS Statistics for Windows, Version 26.0. Armonk, NY: IBM Corp) with P-values significance threshold of 0.05.

Features with more than 10% missing values were entirely omitted. Imputation via the k-nearest neighbor (KNN) algorithm with k=5 and uniform weights was used for features with less than 10% missing values; KNN algorithm imputes every missing value using the mean value from 'k' closest data points found in the training set.

Support Vector Machine Analysis

Support vector Machine (SVM) classifiers were developed based on demographic (2 feature), Laboratory (26 feature) and clinical (9 feature) information (Figure 2).

The SVM binary classification algorithm searches for an optimal hyperplane that separates the data into two classes. For separable classes, the optimal hyperplane maximizes a margin (space that does not contain any observations) surrounding itself, which creates boundaries for the positive and negative classes. For inseparable classes, the objective is the same, but the algorithm imposes a penalty on the length of the margin for every observation that is on the wrong side of its class boundary. The linear SVM score function for a binary classification problem is as below.

$$f(x) = x^t\beta + b$$

Where x is an observation and vector β contains the coefficients of decision the hyperplane. For a particular hyperplane, f(z) is the distance from point z to the hyperplane. The algorithm searches for the maximum margin length, while keeping observations in the positive (y = 1) and negative (y = -1) classes separate. The package used in this study solves the dual formalization using the Lagrange multipliers method to optimize the objective. The dual formalizations of linear SVM for separable classes are as follows:

$$0.5 \sum_{j=1}^n \sum_{k=1}^n \alpha_j \alpha_k y_j y_k x_j^t x_k^t - \sum_{j=1}^n \alpha_j$$

with respect to $\alpha_1 \dots \alpha_n$ subject to

$$\sum_{j=1}^n \alpha_n y_j = 0 \text{ for all } j = 1 : n \text{ and Karush-Kuhn-Tucker (KKT) complementarity conditions.}$$

where the $\alpha_1 \dots \alpha_n$ are the n coefficients of Lagrange multipliers [33]. Among various data mining methods, SVM is well known for its classification discriminative power, especially in cases where sample sizes are small and a relatively large number of features are present (i.e., high-dimensional space) similar to the current data set (number of features = 37, sample number = 242). The hyperparameters of each

classifier were fine-tuned on the training and tuning sets. The best model was evaluated on the test set. For the SVM classifier, linear kernels were used to untangle further interpretations [33] of model's most contributing factors. The algorithm training and cross-validations have been implemented in MATLAB using Statistics and Machine Learning Toolbox™ package [34].

Neighborhood Component Analysis

To investigate the principal features for mortality prediction, Neighborhood component analysis (NCA) was utilized. This non-parametric analysis is an embedded method for selecting features with the goal of maximizing prediction accuracy of regression and classification algorithms. The Statistics and Machine Learning Toolbox in MATLAB[34] version R2019b was used for implementation. functions *fscnca* and *fsrnca* perform NCA feature selection with regularization to learn feature weights for minimization of an objective function that measures the average leave-one-out classification or regression loss over the training data [35].

Receiver Operating Characteristic (ROC) curve

An ROC curve is a graph showing the performance of a classification model at all classification thresholds with illustrating True Positive Rate (TPR) versus False Positive Rate (FPR) as defined below:

$$TPR = \frac{TP}{TP+FN}$$

$$FPR = \frac{FP}{FP+TN}$$

Since TPR is equivalent to and FPR is equal to $1 - specificity$, the ROC graph is sometimes called the sensitivity vs $1 - specificity$ plot. The Area Under Curve(AUC) of the ROC plot is an estimate of the probability that a classifier will rank a randomly chosen positive instance higher than a randomly chosen negative instance. We used AUC of SVM model to compare invasive and non-invasive models. To calculate the optimal operating point of the ROC curve, the algorithm first finds the slope, S, using

$$S = \frac{Cost(P|N) - Cost(N|N)}{Cost(N|P) - Cost(P|P)} * \frac{N}{P}$$

Where the $Cost(N|P)$ is the cost of misclassifying a positive class as a negative class and $Cost(P|N)$ is the cost of misclassifying a negative class as a positive class. P and T are the total instance counts in the positive and negative class, respectively. $P = TP + FN$ and $N = TN + FP$.

Declarations

Author contributions statement

E.K., R.L., H.C. and M.M. conceptualized the idea and initialized the project. E.K. and R.L. supervised the project. H.C. and M.M. collected the data. A.A., A.K., M.R., and Z.J. provided clinical consultation. H.C, M.M., and E.Z. Analyzed the data. A.G. provided analysis consultation. E.K., R.L., H.C., M.M., and E.Z., drafted the manuscript. All authors provided critical review of the manuscript and approved the final draft for publication.

References

1. WHO Coronavirus Disease (COVID-19) Dashboard <https://covid19.who.int>. (accessed: 09.02.2020).
2. Belongia, E. A. & Osterholm, M. *COVID-19 and flu, a perfect storm* 2020.
3. Li, Q. *et al.* Hematological features of persons with COVID-19. *Leukemia*, 1–10 (2020).
4. Quah, , Li, A. & Phua, J. Mortality rates of patients with COVID-19 in the intensive care unit: a systematic review of the emerging literature. *Critical Care* **24**, 1–4 (2020).
5. Cascella, , Rajnik, M., Cuomo, A., Dulebohn, S. C. & Di Napoli, R. in *Statpearls [internet]* (StatPearls Publishing, 2020).
6. Mei, X. *et al.* Artificial intelligence-enabled rapid diagnosis of patients with COVID-19. *Nature Medicine*, 1–5 (2020).
7. Mukherjee, *et al.* A shallow convolutional neural network predicts prognosis of lung cancer patients in multi-institutional computed tomography image datasets. *Nature Machine Intelligence* **2**, 274–282 (2020).
8. Yan, *et al.* An interpretable mortality prediction model for COVID-19 patients. *Nature Machine Intelligence*, 1–6 (2020).
9. Xie, *et al.* Development and external validation of a prognostic multivariable model on admission for hospitalized patients with COVID-19 (2020).
10. Yadaw, S. *et al.* Clinical predictors of COVID-19 mortality. *medRxiv* (2020).
11. Liang, *et al.* Early triage of critically ill COVID-19 patients using deep learning. *Nature communications* **11**, 1–7 (2020).
12. Wu, & McGoogan, J. M. Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China: summary of a report of 72 314 cases from the Chinese Center for Disease Control and Prevention. *Jama* **323**, 1239–1242 (2020).
13. Pan, *et al.* A Predicting Nomogram for Mortality in Patients With COVID-19. *Frontiers in Public Health* **8**, 461 (2020).
14. Hastie, , Tibshirani, R. & Friedman, J. *The elements of statistical learning: data mining, inference, and prediction* (Springer Science & Business Media, 2009).
15. Van Bavel, J. *et al.* Using social and behavioural science to support COVID-19 pandemic response. *Nature Human Behaviour*, 1–12 (2020).
16. Siordia Jr, A. Epidemiology and clinical features of COVID-19: A review of current literature. *Journal of Clinical Virology*, 104357 (2020).

17. Kotas, E. & Medzhitov, R. Homeostasis, inflammation, and disease susceptibility. *Cell* **160**, 816–827 (2015).
18. Graves, T. & Kayal, R. A. Diabetic complications and dysregulated innate immunity. *Frontiers in bioscience: a journal and virtual library* **13**, 1227 (2008).
19. Kose, & Mandiracioglu, A. Fear of blood/injection in healthy and unhealthy adults admitted to a teaching hospital. *International Journal of Clinical Practice* **61**, 453–457 (2007).
20. Plebani, M. Diagnostic errors and laboratory medicine—causes and strategies. *Ejifcc* **26**, 7 (2015).
21. Reeve, J. L. & Twomey, J. Consider laboratory aspects in developing patient prediction models. *Nature Machine Intelligence*, 1–1 (2020).
22. Carroll, A. E. The high costs of unnecessary care. *Jama* **318**, 1748–1749 (2017).
23. Maves, R. C. *et al.* Triage of scarce critical care resources in COVID-19: an implementation guide for regional allocation An expert panel report of the Task Force for Mass Critical Care and the American College of Chest Physicians. *Chest* (2020).
24. Dias, D. & Paulo Silva Cunha, J. Wearable health devices—vital sign monitoring, systems and technologies. *Sensors* **18**, 2414 (2018).
25. Zhou, *et al.* Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *The lancet* (2020).
26. Smits, L. *et al.* Exacerbated innate host response to SARS-CoV in aged non-human primates. *PLoS Pathog* **6**, e1000756 (2010).
27. Opal, M., Girard, T. D. & Ely, E. W. The immunopathogenesis of sepsis in elderly patients. *Clinical infectious diseases* **41**, S504–S512 (2005).
28. Teo, Early detection of silent hypoxia in COVID-19 pneumonia using Smartphone pulse oximetry. *Journal of medical systems* **44**, 1–2 (2020).
29. Xie, J. *et al.* Association between hypoxemia and mortality in patients with COVID-19 in Mayo Clinic Proceedings (2020).
30. Giacobbe, D. R. Clinical interpretation of an interpretable prognostic model for patients with COVID-19. *Nature Machine Intelligence*, 1–1 (2020).
31. Yang, *et al.* COVID-19: immunopathogenesis and Immunotherapeutics. *Signal Transduction and Targeted Therapy* **5**, 1–8 (2020).
32. Connors, M. & Levy, J. H. Thromboinflammation and the hypercoagulability of COVID-19. *Journal of Thrombosis and Haemostasis* (2020).
33. Cristianini, N., Shawe-Taylor, , *et al.* *An introduction to support vector machines and other kernel-based learning methods* (Cambridge university press, 2000).
34. *Statistics and Machine Learning Toolbox™ User's Guide R2019b*. 2019.
35. Yang,, Wang, K. & Zuo, W. Neighborhood Component Feature Selection for High-Dimensional Data. *JCP* **7**, 161–168 (2012).

Tables

Due to technical limitations, tables 1-2 are only available as a download in the supplemental files section.

Figures

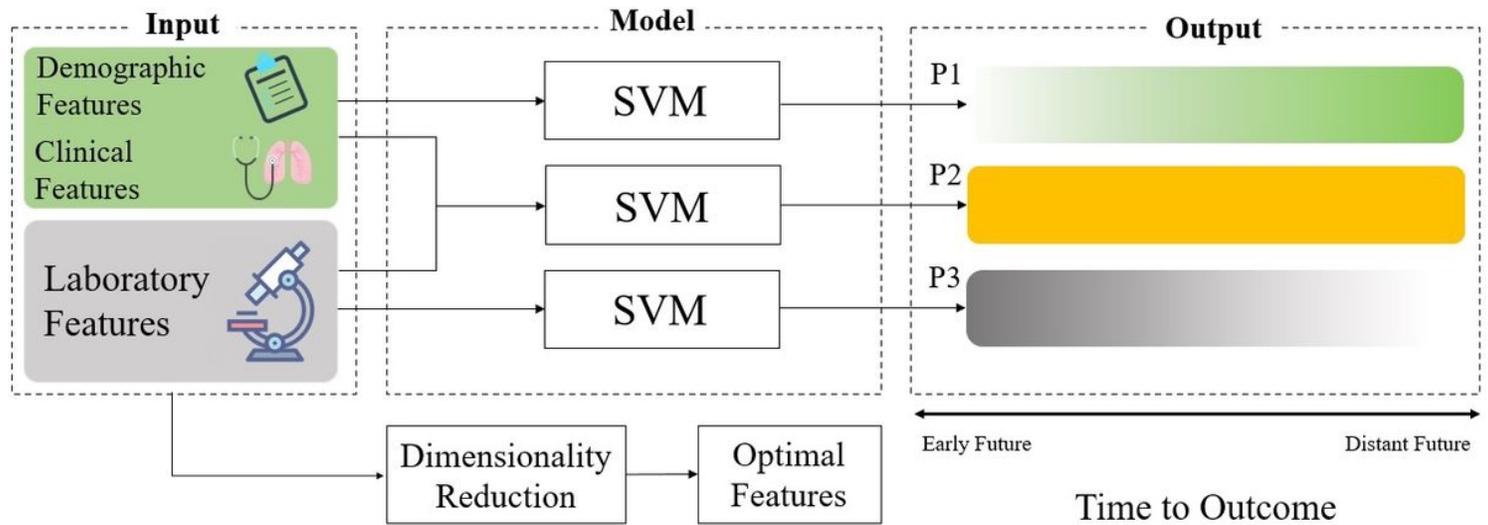


Figure 1

Illustration of modelling framework. Three machine learning models were developed using SVM framework with three input groups; invasive, non-invasive, and their combination. Invasive group comprises laboratory results. Non-invasive features comprises patient clinical and demographic data. Joint group comprises the combination of invasive and non-invasive features. P1, P2, and P3 represent the prediction performance provided by the non-invasive, joint, and invasive models, respectively. The non-invasive model displayed good prediction performance in farther future (P1) whereas invasive model showed good prediction performance for near future (P3). In a parallel dimensionality reduction analysis, the NCA algorithm was utilized to determined the minimal number of features that can approximately capture the full model accuracy

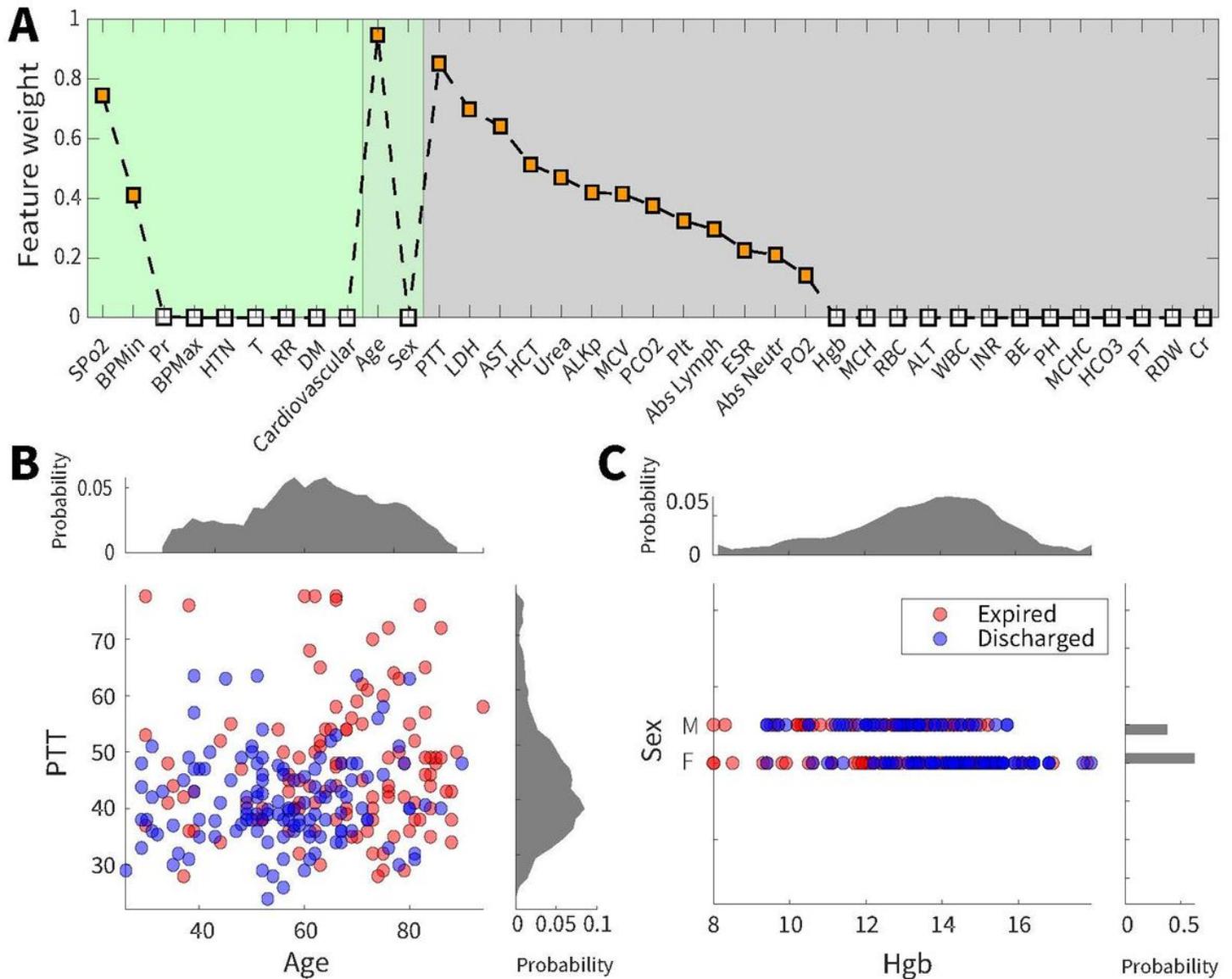


Figure 2

Contribution of demographic, clinical, and laboratory single features to mortality prediction (A) The NCA implemented to calculate feature contributions in mortality prediction. Features are sorted based on contribution weights and category. (B) is a favorable feature space (SPO2 and PCO2) where many data points could be distinguished and (C), in contrast, demonstrates unfavorable feature space (Sex and Hgb).

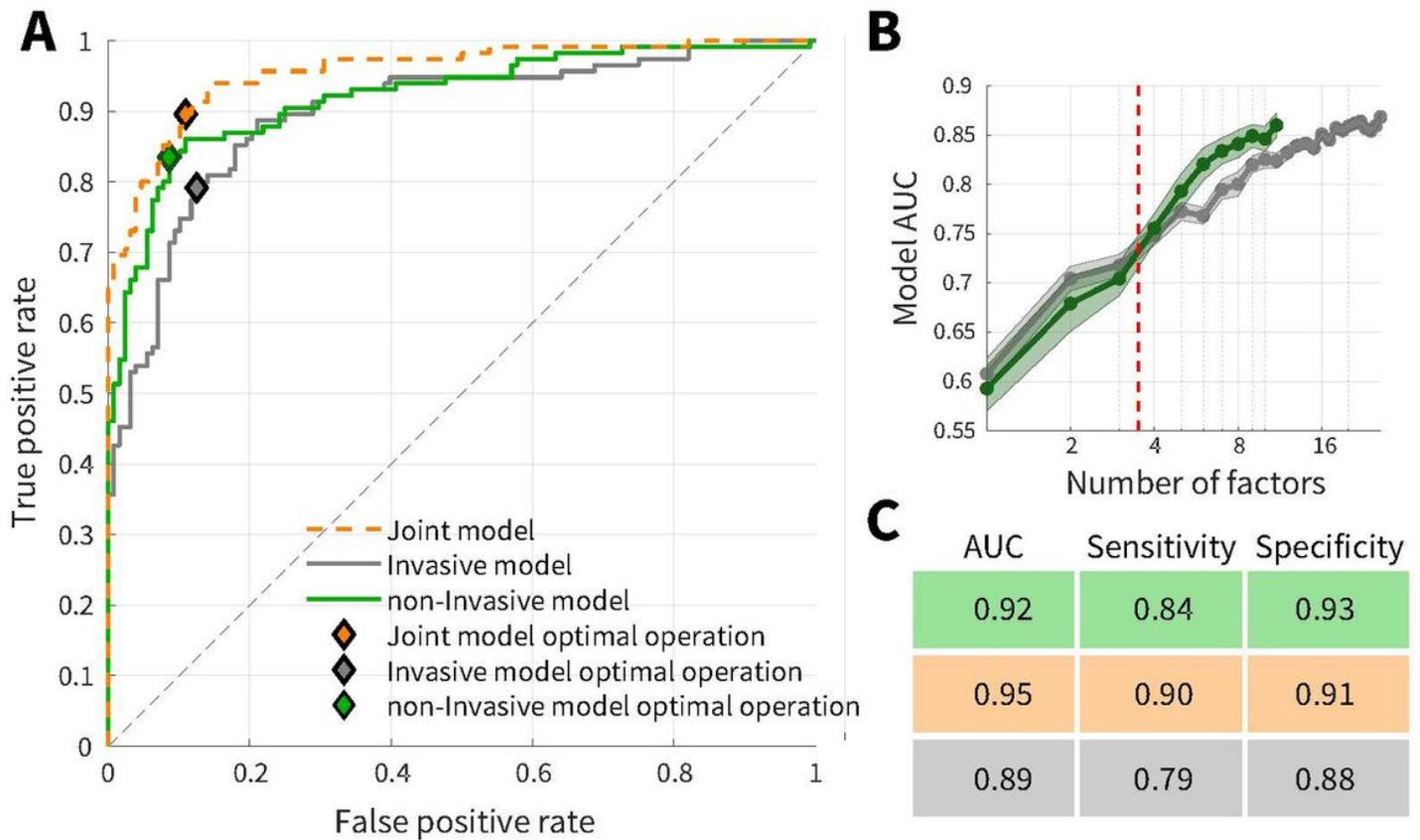


Figure 3

Prediction performance of machine learning models. (A) ROC curve of joint, invasive, and non-invasive prediction models. (B) Comparison of model AUC between invasive and non-invasive models across feature space size. This plot indicates that non-invasive models with 4 or higher number of input features provide better performance compared with invasive models with same size. (C) Performance of joint (orange), invasive (gray), and non-invasive (green) models

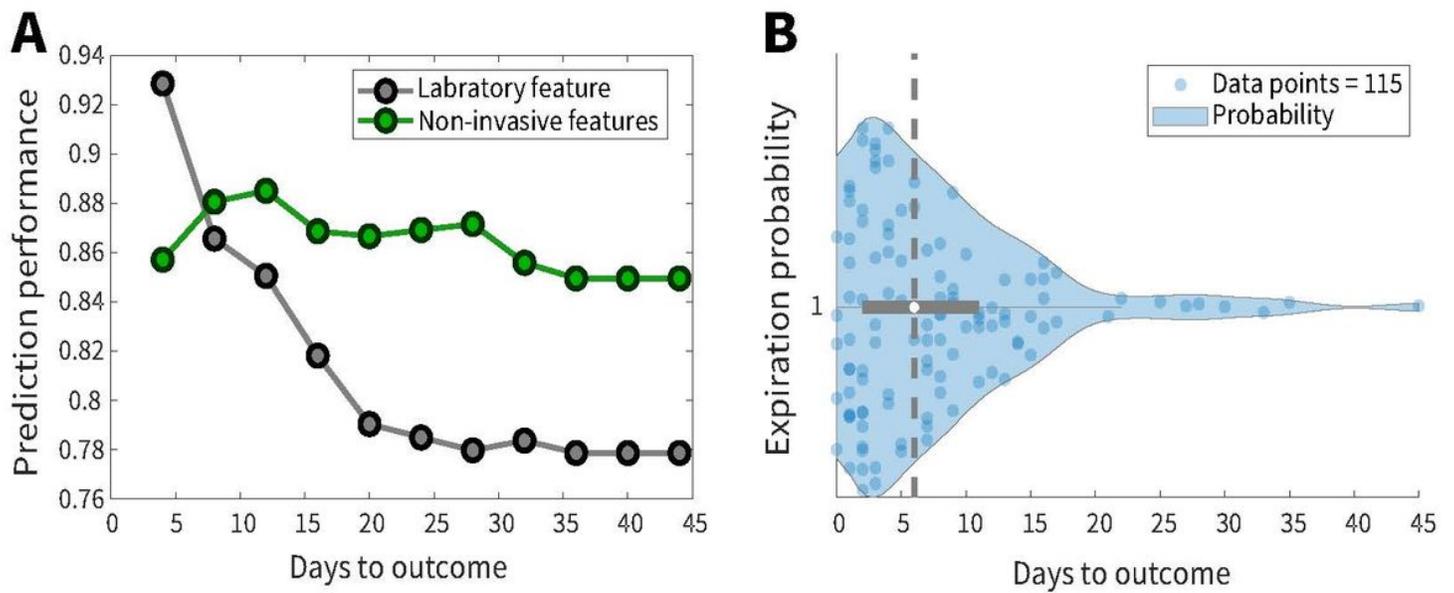


Figure 4

Temporal forecasting range of invasive and non-invasive models. (A) Prediction performance of invasive and non-invasive models across expiration temporal spectrum. Days to outcome was defined as the number of days between patient admission and expiration. While invasive features were better predictors for imminent future, they were outperformed by non-invasive features over larger expiration intervals. (B) Temporal distribution of patient expiration intervals. Vertical dashed line corresponds to the median expiration interval which was 6 days after admission.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [FigureS1.jpg](#)
- [NatureCommunicationsTables20200909.pdf](#)