Workshop

Nicolò Casano, Silvano Junior Santini, Pierpaolo Vittorini, Gaia Sinatti, Paolo Carducci, Claudio Maria Mastroianni, Maria Rosa Ciardi, Patrizia Pasculli, Emiliano Petrucci, Franco Marinangeli and Clara Balsano*

Application of machine learning approach in emergency department to support clinical decision making for SARS-CoV-2 infected patients

https://doi.org/10.1515/jib-2022-0047

Received September 20, 2022; accepted February 8, 2023; published online March 6, 2023

Abstract: To support physicians in clinical decision process on patients affected by Coronavirus Disease 2019 (COVID-19) in areas with a low vaccination rate, we devised and evaluated the performances of several machine learning (ML) classifiers fed with readily available clinical and laboratory data. Our observational retrospective study collected data from a cohort of 779 COVID-19 patients presenting to three hospitals of the Lazio-Abruzzo area (Italy). Based on a different selection of clinical and respiratory (ROX index and PaO2/FiO2 ratio) variables, we devised an AI-driven tool to predict safe discharge from ED, disease severity and mortality during hospitalization. To predict safe discharge our best classifier is an RF integrated with ROX index that reached AUC of 0.96. To predict disease severity the best classifier was an RF integrated with ROX index that reached an AUC of 0.91. For mortality prediction the best classifier was an RF integrated with ROX index, that reached an AUC of 0.91. The results obtained thanks to our algorithms are consistent with the scientific literature an accomplish significant performances to forecast safe discharge from ED and severe clinical course of COVID-19.

Keywords: COVID-19; emergency medicine; machine learning; ROX index; SARS-CoV-2.

1 Introduction

The pandemic caused by severe respiratory acute syndrome coronavirus 2 (SARS-CoV-2) was declared global emergency on 11th March 2020, by the World Health Organization (WHO) [1]. Since then, the development and

***Corresponding author: Clara Balsano**, School of Emergency Medicine, Interdisciplinary BioMedical group on Artificial Intelligence, IBMAI, Department MeSVA, University of L'Aquila, Piazza S. Salvatore Tommasi 1, 67100, Coppito, L'Aquila, Italy; and Francesco Balsano Foundation, Via Giovanni Battista Martini 6, 00198, Rome, Italy, E-mail: clara.balsano@univaq.it

Nicolò Casano, **Pierpaolo Vittorini and Paolo Carducci**, School of Emergency Medicine, Interdisciplinary BioMedical group on Artificial Intelligence, IBMAI, Department MeSVA, University of L'Aquila, L'Aquila, Italy. https://orcid.org/0000-0002-6975-8958 (P. Vittorini)

Silvano Junior Santini and Gaia Sinatti, School of Emergency Medicine, Interdisciplinary BioMedical group on Artificial Intelligence, IBMAI, Department MeSVA, University of L'Aquila, L'Aquila, Italy; and Francesco Balsano Foundation, Via Giovanni Battista Martini 6, 00198, Rome, Italy

Claudio Maria Mastroianni, **Maria Rosa Ciardi and Patrizia Pasculli**, Department of Public Health and Infectious Diseases, "Sapienza" University of Rome, Policlinico Umberto I Hospital, Rome, Italy

Emiliano Petrucci and Franco Marinangeli, Department of Anesthesiology, Intensive Care and Pain Treatment, University of L'Aquila, L'Aquila, Italy

Open Access. © 2023 the author(s), published by De Gruyter. This work is licensed under the Creative Commons Attribution 4.0 International License.

distribution of Coronavirus Disease 2019 (COVID-19) mRNA vaccines, and the efforts made to select more effective antiviral and immunomodulatory therapies have significantly reduced the incidence of severe disease and mortality among patients [2–4].

Although these progresses have changed the face of pandemic, some concerns remain about threats related to SARS-CoV-2 [5].

By example, some countries have not adopted a common global response to SARS-CoV-2 infection and vaccination inequities are manifest [6]. As of July 2022, the 85% of vaccine doses have been administrated in high income countries; in low-income countries the accessibility to vaccination is the principal cause of unvaccinated patients, especially in the African continent, where for some countries the fully vaccinated protocol was administrated to the 10–20% of population [7–10]. This scenario confirms that recurrence of viral outbreak remains a concrete challenge, especially in countries where the doses of vaccines administered are inadequate.

An early clinical prediction of clinical outcome could be beneficial for the patient's management and for the optimal allocation of healthcare resources [11, 12]. As we previously experienced, the clinical burden was higher in non-metropolitan areas than metropolitan ones, where the hospitals and medical services were less prepared for a viral outbreak [13]. Artificial intelligence and related tools could overcome these issues by supporting physicians in clinical routine. Accordingly, several clinical indexes, machine learning (ML) models and apps have been demonstrated to be a powerful approach for predicting hospitalization, disease severity and mortality in ED [14–18].

We performed an overview of several state-of-the-art ML classifiers that, starting from routine clinical and laboratory analyses, could support physician's decision making in achieving a safe discharge, as well as in forecasting disease severity and mortality risk of SARS-CoV-2 infected patients. All the ML classifiers, used in the study, were compared with the indexes to assess the risk of severe disease or death, commonly used in ED.

The work presented in this paper is part of a larger research project whose aim was to design, develop and test a decision support system, called C19DSS (COVID-19 Decision Support System), based on ML predictors, that should assist physicians in clinical decision making [19].

2 Materials and methods

2.1 Study design and sample

This study retrospectively analyzed a cohort of SARS-CoV-2 infected patients presenting to ED, to collect a dataset for devising the ML models. Biochemical and clinical data of 779 adults (age > 18 years), affected by SARS-CoV-2, admitted during the first two waves to three hospitals of Lazio-Abruzzo area, were collected and analyzed. Missing data were managed by automatic imputation as described in the section below.

Demographic, clinical, and biochemical variables were collected by a physician, during the first presentation at each ED.

The clinical outcomes were safe discharge, disease severity at seventh day of hospitalization and mortality during hospitalization. Data were collected at three different times: at ED presentation (T0), at the seventh day of hospitalization (T7) and at hospital discharge (TD). The disease severity was interpreted by following the interim guidance released by the WHO [20]. WHO ordered the severity of the infectious disease in "mild" (WHO score 0), "moderate" for patients who no required supplemental oxygen (WHO score 1), and "severe" for all patients who needed non-invasive ventilation (NIV) or high-flow oxygen requirement or continuous positive airway pressure (CPAP) (WHO score 2), and "critical" for patients who needed invasive ventilation or multi-organ support (WHO score ≥3). The safe discharge was a composite endpoint considering for all enrolled patients: PaO_2/FiO_2 ratio > 274, PaO_2 > 65 mmHg, no oxygen supply during hospitalization, length of hospitalization < 10 days and no adverse events during hospitalization (endotracheal intubation, disease progression to severe or critical disease, admission to high dependency unit, intensive care unit, mortality). The cut-off of PaO_2/FiO_2 ratio > 274 was chosen considering the results reported in our previous manuscript [21]. The safe discharge outcome reported in the dataset was then revised, so to represent the correct decision that should have been taken at admission time. The mortality risk and the severity disease risk were retrospectively evaluated, applying respectively the 4C Mortality Score for COVID-19 (4C MS) and the quick COVID-19 Severity Index (qCSI) [22, 23].

The study was approved by the Internal Review Board of our Institution (IRB 06/2021, protocol number 28,958) and registered on clinicaltrial.gov, protocol number NCT04825301. The study was conformed to the Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD) statement and should be considered a prediction model study type 2a [24].

2.2 Missing data imputation

To devise the models in case of missing data, the dataset was pre-processed through missing data imputation system [25, 26]. In detail, predictive mean matching for numeric data and logistic regression imputation for categorical data (with 2 levels) was used [27].

2.3 Predictors

To devise the ML models, both continuous and categorical variables were used. Heart rate (HR), respiratory rate-oxygenation (ROX) index, alanine aminotransferase (ALT), aspartate aminotransferase (AST), creatinine, C-reactive protein (CRP), lactic dehydrogenase (LDH), platelet count (PLT), white blood cells count (WBC), neutrophil count (NEU), lymphocyte count (LIN), eosinophil count (EOS), basophil count (BAS), neutrophil lymphocyte ratio (NLR), platelet lymphocyte ratio (PLR), international normalized ratio (INR), D-dimer (DD), standard blood gas analysis variables and PaO_2/FiO_2 ratio have been included as continuous variables. Sex at birth, fever (or not) and cough (or not) have been considered as categorical variables.

2.4 Outcomes

Primary outcome was the prediction (i.e., safe/risk) of safe discharge; secondary outcomes were the prediction of disease severity at seventh day (i.e., severe/mild) and mortality during hospitalization (i.e., present/absent).

2.5 Statistical analysis and ML models

The descriptive analyses consist of frequency tables, medians and IQRs. Data reported in Supplementary Table 1 have been analyzed using one way ANOVA with post-hoc Fisher LSD tests for multiple comparison.

Based on a different selection of predictors, several models with complete set of predictors and with set of predictors were devised, details are reported in Supplementary Table 3.

The ML task for all outcomes is binary classification. The following different models were tried from the available data: decision tree (DT), random forest (RF) and gradient boosting machines (GBM) [28–30]. For all models and outcomes, the dataset was split into train and test (with 75% of data going for training, 25% for testing), used 10-fold cross-validation, tuned each classifier according to its specific hyperparameters, calculated the confusion matrix (with the most useful associated statistics, i.e., accuracy, Cohen's K, sensitivity, specificity) and the ROC curve (with the related AUC) [31]. To consider data imbalance, we also adopted RUSBoost as hybrid sampling/boosting algorithm [32].

In addition, to improve the explainability of the models, the variables importance was calculated and presented as graphs [33].

To compare the ML models (for the disease severity and mortality risk) with the standard clinical practice, the qCSI and 4C MS indexes was dichotomized. The confusion matrices (with the same associated statistics of the ML models) were calculated and finally, the results were compared with the predictions of the corresponding ML models, i.e., severity and mortality, respectively.

All statistical analyses and ML models discussed in the paper were calculated with RStudio 2022.07.2–576 running R4.2.1 for Linux, using the "caret" and "mice" packages [27, 34]. Due to the restrictions decided by the Internal Review Board, we are unable to publicly share the study data.

3 Results

3.1 Sample description

The final dataset included 779 patients; demographic and clinical characteristics and principal biochemical markers for the study population are reported in Supplementary Table 1. Data record was not complete for 283 patients, whereas 496 patients have got a complete data record. The study population presented a median age of 65 years (IQR 55–75), 41.3% were female, more than three comorbidities were present in 19% of the cases. The median value of CRP was 4.84 mg/dL (IQR 1.58–10.89), of LDH was 302 IU/L (IQR 227–389.5), of platelet was 202 10^3/mcl (IQR 160–265), of WBC was 6.27 10^9/L (IQR 4.60–9.24), of neutrophil count was normal 4.55 10^9/L/L (IQR 3.11–7.60), of lymphocytes count was 0.92 10^9/L/L (IQR 0.64–1.33). D-Dimer was tested in D-Dimer Unit (DDU), its median value was 0.81 mcg/mL DDU (IQR 0.47–1.50). Respiratory parameter as PaO_2 had

a median value of 70.2 mmHg (IQR 59–83), the PaO_2/FiO_2 ratio 308 (IQR 223–365) and the ROX index 20.4 (IQR 13–20.9). Patients with severe disease or deceased were elder and characterized by higher values of CRP, LDH, D-Dimer, and progressive lower values for the PaO_2/FiO_2 ratio and ROX index, compared to the safe discharge population.

At presentation in ED, 72.6% of patients reported fever and 54.1% referred cough; the clinical spectrum of disease of the whole cohort is outlined in Supplementary Table 2. At presentation in ED, SARS-CoV-2 patients presented mild (26.3%), moderate (38.7%), severe (28.6%), critical (6.4%) symptoms. In Supplementary Table 2 are reported the distribution of safe discharge, severe disease and died population considering the WHO interim guidance (see material and methods).

3.2 Machine learning models

As previously described, based on a different selection of predictors we devised several models, the results of all the prediction models developed are reported in Supplementary in Table 4.

Tables 1a and 1b report the main statistics as coming from the confusion matrices for the best classifiers selected: the ROX index models and the PaO_2/FiO_2 ratio models. Figure 1 instead contains all ROC curves with the related AUCs, for the best classifiers (in terms of AUC), for the different outcomes. Finally, Figure 2 reports the importance of the first ten variables, ranked by importance, for the different outcomes, for the best classifiers (in terms of AUC).

PaO_2/FiO_2 ratio models. For safe discharge, RF is the most accurate in almost all the statistics, except for sensitivity, where GBM performs slightly better. The results for disease severity showed that RF performed with the best AUC; conversely, GBM has slightly superior sensitivity and accuracy. Regarding mortality during hospitalization, GBM reached the best AUC and specificity, on the other hand, RF performed the best accuracy and sensitivity (Table 1a and Figure 1).

ROX index models. For safe discharge, GBM has the best AUC. On the other hand, RF demonstrated the highest sensitivity and accuracy. The results for disease severity showed that RF is the most accurate model, with the highest specificity and the best AUC. Conversely, GBM has the best accuracy. In connection with mortality during hospitalization, RF is the best classifier in terms of all statistics, except for specificity, where GBM performs slightly better (Table 1b and Figure 1).

Table 1a: PaO_2/FiO_2 ratio models: main statistics for all outcomes and classifiers.

	PaO_2/FiO_2 ratio models				**PaO_2/FiO_2 ratio, RUSBoost models**			
	Acc.	**Sens.**	**Spec.**	**AUC**	**Acc.**	**Sens.**	**Spec.**	**AUC**
				Safe discharge				
DT	0.824	0.842	0.770	0.858	0.837	0.811	0.954	0.927
RF	**0.829**	0.869	**0.780**	**0.894**	0.837	0.802	**1.00**	**0.932**
GBM	0.824	**0.876**	0.666	0.882	**0.845**	**0.831**	0.909	0.908
				Disease severity				
DT	0.742	0.732	**0.752**	0.766	0.699	**0.729**	0.653	0.773
RF	0.757	0.783	0.732	**0.832**	**0.739**	**0.729**	**0.755**	**0.815**
GBM	**0.762**	**0.804**	0.721	0.827	0.707	0.689	0.734	0.773
				Mortality				
DT	0.840	0.944	0.290	0.689	0.772	0.818	0.583	0.759
RF	**0.876**	**0.969**	0.387	0.842	0.788	0.808	**0.708**	0.856
GBM	0.860	0.944	**0.419**	**0.844**	**0.804**	**0.838**	0.666	**0.864**

The best result is reported as bold value.

Table 1b: ROX index models: main statistics for all outcomes and classifiers.

	ROX index models				ROX index, RUSboost models			
	Acc.	**Sens.**	**Spec.**	**AUC**	**Acc.**	**Sens.**	**Spec.**	**AUC**
				Safe discharge				
DT	**0.928**	0.939	**0.875**	0.778	**0.938**	**0.939**	0.937	0.960
RF	**0.928**	**0.963**	0.750	0.879	0.928	0.914	**1.00**	**0.962**
GBM	0.898	0.951	0.625	**0.898**	0.898	0.890	0.937	0.955
				Disease severity				
DT	0.826	**0.857**	0.809	0.709	0.816	0.809	0.828	0.877
RF	0.816	0.742	**0.857**	**0.806**	**0.846**	**0.841**	0.857	**0.911**
GBM	**0.836**	0.800	**0.857**	0.800	0.795	0.730	**0.914**	0.910
				Mortality				
DT	0.827	**0.974**	0.318	0.735	0.787	0.779	**0.818**	0.864
RF	**0.828**	**0.974**	0.318	**0.860**	**0.858**	**0.883**	0.772	**0.915**
GBM	**0.828**	0.935	**0.454**	0.844	0.777	0.792	0.727	0.878

3.3 Comparison of ML models with standard clinical practice

Table 2 contains the main statistics of the best classifiers and the indexes commonly used in clinical practice, for both disease severity and mortality. Figure 3 depicts the comparison in terms of ROC curves and AUCs. As mentioned earlier, given that the indexes could be computed only for the complete cases, the comparison is reported with the models devised from the same subset of observations. As previously mentioned, the main statistics for complete cases are reported in Supplementary Table 4. The results show that our models for predicting mortality outperforms the 4C MS in all statistics. On the other hand, with respect to qCSI, our models for predicting disease severity showed higher accuracy, specificity and AUC, but a smaller sensitivity.

3.4 External validation

The best prediction models were selected and integrated into a system; further details are described in the paragraph below (see **COVID-19 Decision Support System**). During the fourth wave, the system was used by the physicians of different institutions and was applied to patients with infection by SARS-CoV-2. So far, the validation cohort is made up of 57 patients. The mortality risk prediction model showed an accuracy of 0.94 for both GBM and RF models, whereas the models about safe discharge returned an accuracy of 0.70 and 0.60 (RF and GBM respectively).

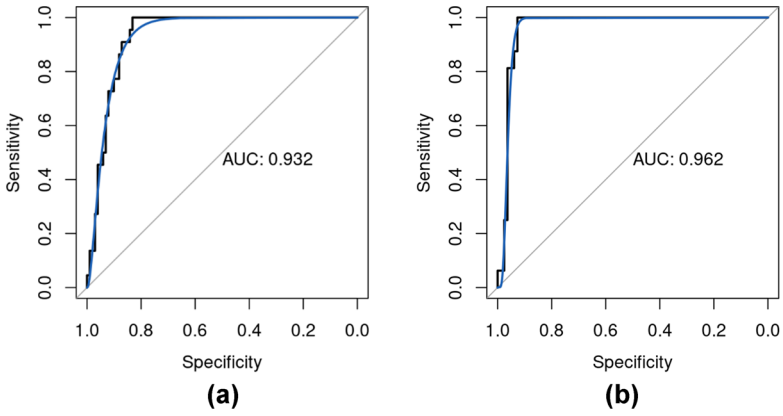
Therefore, as expected, the system confirms its high sensitivity, but a lower overall accuracy [34, 35].

3.5 COVID-19 Decision Support System (C19DSS)

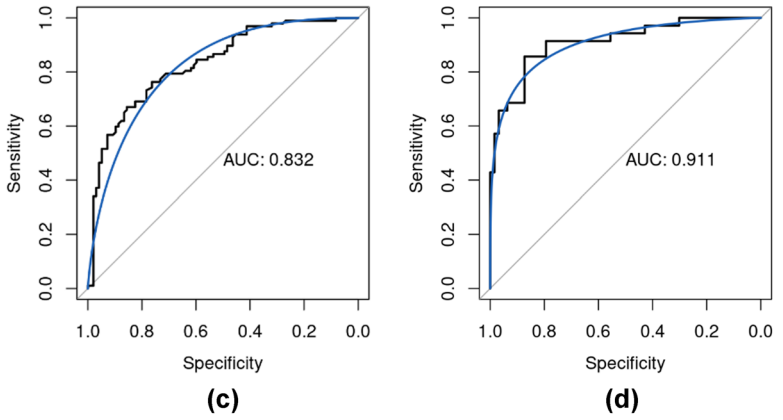
We developed the COVID-19 Decision Support System (C19DSS) to enable physicians to effectively use the previous models. The system is made up of a smartphone app used by clinicians, and a server that provides the "intelligence" to the app.

In short, the app is made up of four activities (see Supplementary Figure 1). The first activity is the dashboard, that summarizes the database content and the server connection status. The second activity contains the patient list, plus a set of filters and a button to add a new patient. The third screen shows how to enter the laboratory/clinical data of a new patient, and the button to request the classification to the server. The fourth screen

A. SAFE DISCHARGE



B. DISEASE SEVERITY



C. MORTALITY

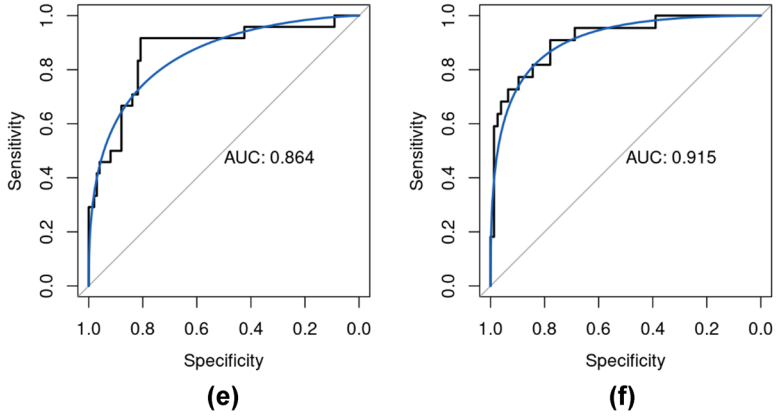


Figure 1: ROC curves of the best models (in terms of AUC) for (A) safe discharge: (a) PaO_2/FiO_2 ratio RUSBoost model – RF, (b) ROX index RUSBoost model – RF; (B) disease severity: (c) PaO_2/FiO_2 ratio model – RF, (d) ROX index RUSBoost model – RF; (C) mortality: (e) PaO_2/FiO_2 ratio RUSBoost model – GBM, (f) ROX index RUSBoost model – RF.

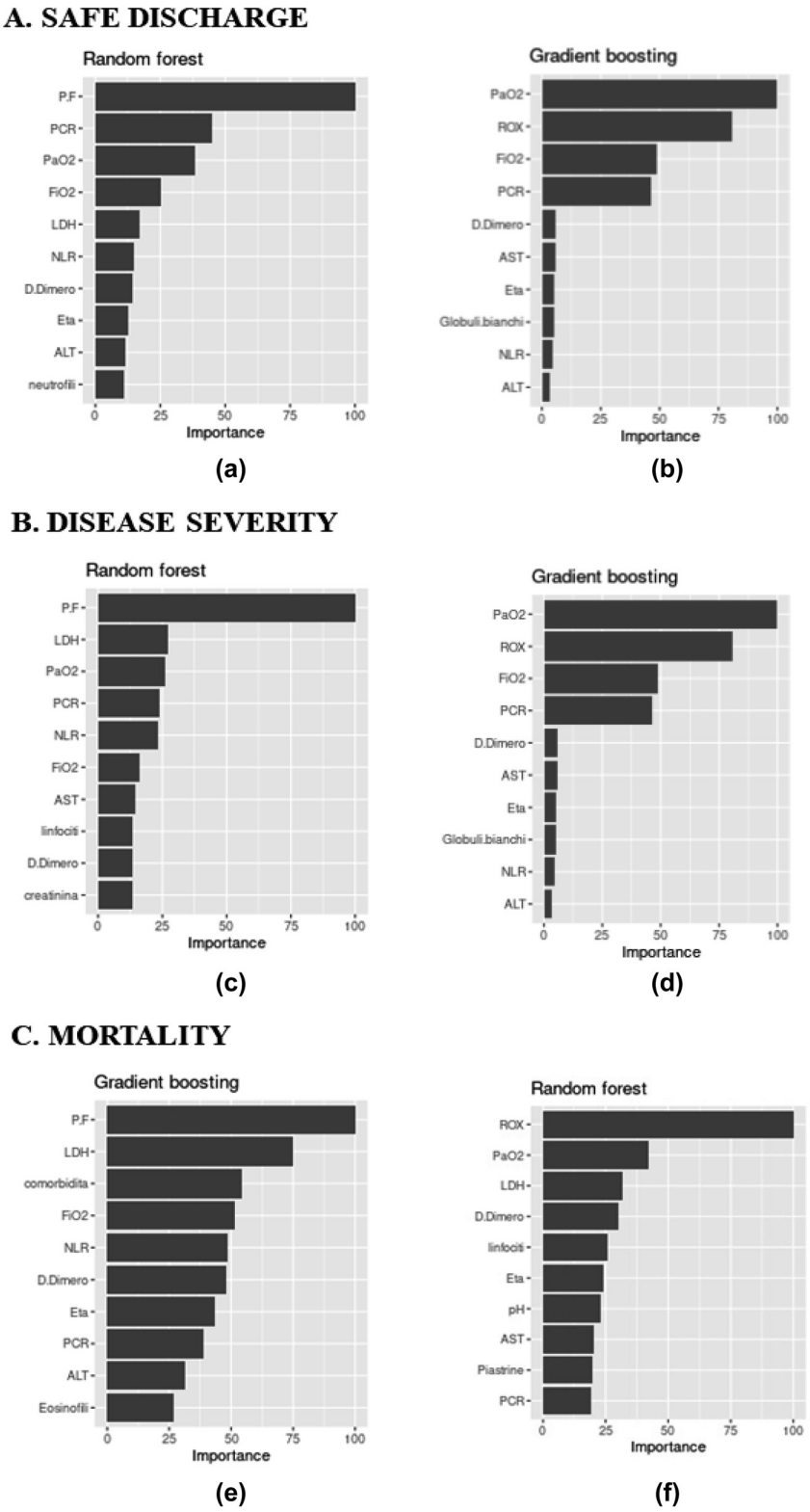


Figure 2: Variables ranked by importance for the different outcomes for (A) safe discharge: (a) PaO_2/FiO_2 ratio model – RF, (b) ROX index model – GBM; (B) disease severity: (c) PaO_2/FiO_2 ratio model – RF, (d) ROX index model – RF; (C) mortality: (e) PaO_2/FiO_2 ratio model – GBM, (f) ROX index model – RF.

Table 2: Main statistics of the standard indexes for disease severity and mortality, regarding our best classifiers.

	Acc.	Sens.	Spec.	AUC
Disease severity				
PaO_2/FiO_2 ratio model – GBM	**0.846**	0.600	**0.908**	0.893
ROX index model – GBM	0.836	0.800	0.857	**0.917**
qCSI	0.805	**0.958**	0.535	0.770
Mortality				
PaO_2/FiO_2 ratio model - RF	**0.878**	0.960	**0.542**	0.866
ROX index model – RF	0.828	**0.974**	0.318	**0.906**
4C MS	0.618	0.958	0.535	0.865

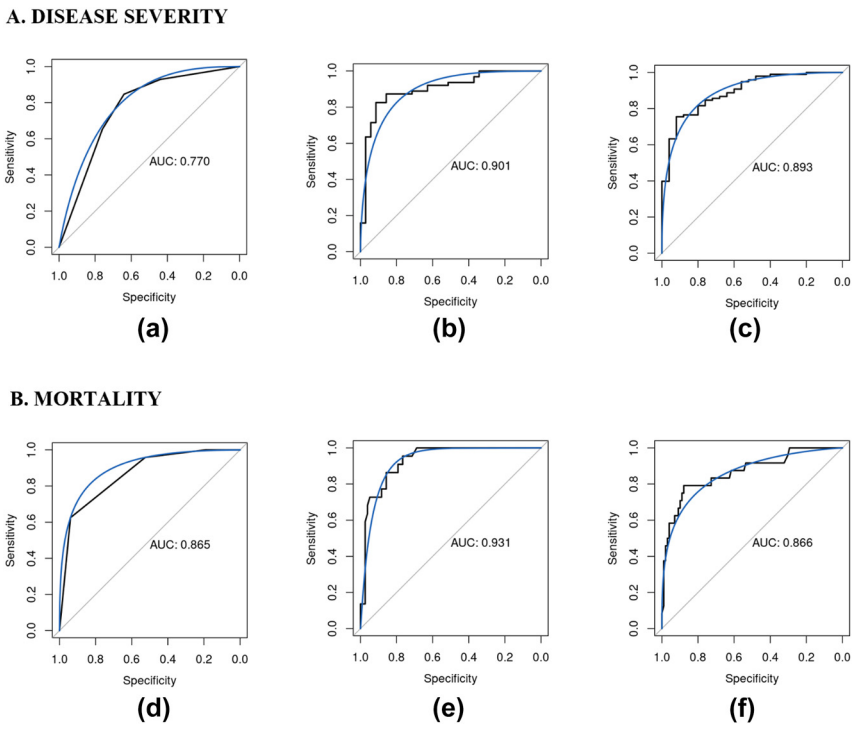


Figure 3: ROC curves and AUCs for (A) disease severity: (a) quick Covid Severity Index (qCSI), (b) ROX index model – GBM, (c) PaO_2/FiO_2 ratio model – GBM; (B) Mortality: (d) 4C Mortality Score (4C MS), (e) ROX index model – RT, (f) PaO_2/FiO_2 ratio model – RF.

depicts how to edit the patient data, as well as request the classification to the server or delete the patient from the local database.

When the classification process is activated, the app opens an encrypted connection to the server, sends the laboratory data and the ID of the patient (so, no personal data is communicated over the network) to the classification service. Then, the server applies the correct model, depending on the request and on the received data, stores the received data for further analyses, and returns the classification results to the app. The activities briefly described above were designed, developed, tested and redesigned following the UCD methodology, i.e., physicians were involved from the very beginning phases of the design, and then we adapted and improved the design/implementation according to consecutive cycles of usability tests. Physicians reported that the application was easy ed intuitive to use; the process of data entry and classification did not hamper the normal ED work

routine. More details on the system, its architecture and on the usability results can be found in a further papers of our research group [19].

4 Discussion

This study was finalized to devise ML methods, fed by routine clinical and laboratory analyses, to be used by physicians to manage safe discharge or hospitalization of unvaccinated COVID-19 patients. In particular, the models aim at predicting three different clinical outcomes: safe discharge, severe disease, at the seventh day after medical presentation, and mortality during hospitalization.

To compare our results with the literature, we proceeded as follows. First, we selected the highest accurate models, coming from peer-reviewed papers, that make use only of clinical data for the classification of the disease without the support of CT or X-ray images [36–39]. The best model found in scientific literature for hospital admission was an RF model that reached an AUC = 0.82. In our research, the ROX index and PaO_2/FiO_2 ratio models used for the prediction of safe discharge reached an AUC of 0.962 and 0.932, respectively. For severity prediction a CatBoost was the best model and reached an AUC equal to 0.84, versus the 0.911 of ROX index model and 0.832 of PaO_2/FiO_2 ratio model. For mortality prediction, the best classifier was a Support Vector Machines (SVM) model that reached an AUC of 0.91, versus 0.864 of ROX index (RF) model and 0.915 of PaO_2/FiO_2 ratio (GBM) model [38, 40, 41]. As can be seen, the RUSboost approach used to consider data imbalance returned the best results, with the only exception of predicting disease severity, in the case of PaO_2/FiO_2 ratio, where a standard RF model achieved better performances.

To our knowledge this is the first study that integrated the ROX index into a machine learning model for clinical prediction [42, 43]. Compared to our PaO_2/FiO_2 ratio model the ROX index model was equal and even slightly superior in all the statistics, and it reached the best performance in the safe discharge model (Figure 1). Although the PaO_2/FiO_2 ratio is the most validated and widely used index for respiratory decompensation, the ROX index is simple to use because it needs only readily available clinical data (RR, SpO_2 and FiO_2) to be calculated [44, 45]. For these properties this simpler model can better help to quickly prioritize patients in the pre-hospital scene and in all the area with limited health resources. Finally, with respect to the standard clinical practice, our models outperform the 4C MS and qCSI indexes [22, 23].

With respect to other models available in the literature, ours are tailored to the clinical and laboratory data that are routinely collected at the first medical presentation in ED.

As expected, the performance in the external validation demonstrated a lower accuracy for both mortality and safe discharge prediction. The reason can be twofold. First, the external validation was conducted on a smaller dataset. Second, the patients of the external validation were enrolled on a second wave of SARS-CoV-2 pandemic respect to those of the internal validation. Accordingly, the evolution and changes of dominant SARS-CoV-2 variants, between the different waves, could have taken a place in this reduced accuracy [46].

5 C19DSS future development

The C19DSS quickly predicted the prognosis in unvaccinated patients with pneumonia due to SARS-CoV-2 infection and, to the best of our knowledge, is the first attempt to provide an accurate prediction of safe discharge from ED. As future work, given the rapid dynamic of the pandemic, our AI-driven prediction models need to be updated to maintain high performance. To this aim, we have programmed to update the models, so that they will take into account the patient's immunization status against SARS-Cov-2 infection, so to improve prediction also in vaccinated, partially vaccinated and immunocompromised patients.

6 Limitations

Besides the limited size of the dataset and the constraint of using only routine clinical and laboratory data to devise the models, the performances of our models are in line and even better than the best prediction models in scientific literature. Furthermore, the prediction models are devised from a cohort of unvaccinated patients, and therefore its applicability is limited to countries with limited access to vaccines. We look forward to subsequent large sample studies and external validations in undeveloped countries.

7 Conclusions

We proposed a novel AI-driven prediction model integrated with the ROX index to forecast safe discharge and mortality in unvaccinated COVID-19 patients.

We conceive that, in the next future, integration of prediction models in emergency care routine could improve healthcare assistance and represent an opportunity to make the people safer.

Author contributions: Clara Balsano (CB), Nicolò Casano (NC), Silvano Junior Santini (SJS), and Pierpaolo Vittorini (PV) significantly contributed to the conception and design of the study. CB, NC, SJS, and PV analyzed the data. All authors were involved in data interpretation and manuscript revision and approved the final version submitted for publication.
Research funding: This study was supported by the Francesco Balsano Foundation (FFB grant 2020/2021). The Internal Review Board of our Institution approved the study (IRB 06/2021, protocol number 28958) and it does not involve any ethical issues.
Conflict of interest statement: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. World Health Organization. WHO Coronavirus (COVID-19) dashboard; 2022. Available from: https://covid19.who.int [Accessed 16 Jul 2022].
2. Beigel JH, Tomashek KM, Dodd LE, Mehta AK, Zingman BS, Kalil AC, et al. Remdesivir for the treatment of covid-19 — final report. N Engl J Med 2020;383:1813—26.
3. Moghadas SM, Vilches TN, Zhang K, Wells CR, Shoukat A, Singer BH, et al. The impact of vaccination on coronavirus disease 2019 (COVID-19) outbreaks in the United States. Clin Infect Dis 2021;73:2257—64.
4. Pinna SM, Lupia T, Scabini S, Vita D, De Benedetto I, Gaviraghi A, et al. Monoclonal antibodies for the treatment of COVID-19 patients: an umbrella to overcome the storm? Int Immunopharm 2021;101:108200.
5. Skegg D, Gluckman P, Boulton G, Hackmann H, Karim SSA, Piot P, et al. Future scenarios for the COVID-19 pandemic. Lancet 2021;397:777—8.
6. Mathieu E, Ritchie H, Ortiz-Ospina E, Roser M, Hasell J, Appel C, et al. A global database of COVID-19 vaccinations. Nat Human Behav 2021;5:947—53.
7. Asundi A, O'Leary C, Bhadelia N. Global COVID-19 vaccine inequity: the scope, the impact, and the challenges. Cell Host Microbe 2021;29:1036—9.
8. Trogen B, Pirofski L anne. Understanding vaccine hesitancy in COVID-19. Nat Med 2021;2:498—501.
9. Mathieu E, Ritchie H, Rodés-Guirao L, Appel C, Giattino C, Hasell J, et al. Coronavirus pandemic (COVID-19) [Online Resource]: Our World In Data. Available from: https://ourworldindata.org/coronavirus [Accessed 23 Jul 2022].
10. Africa Center for Disease Control and Prevention. COVID-19 vaccination. Africa CDC; 2022. Available from: https://africacdc.org/covid-19-vaccination/ [Accessed 22 Aug 2022].
11. Soria A, Galimberti S, Lapadula G, Visco F, Ardini A, Valsecchi MG, et al. The high volume of patients admitted during the SARS-CoV-2 pandemic has an independent harmful impact on in-hospital mortality from COVID-19. PLoS One 2021;16:e0246170.
12. Paganini M, Conti A, Weinstein E, Della Corte F, Ragazzoni L. Translating COVID-19 pandemic surge theory to practice in the emergency department: how to expand structure. Disaster Med Public Health Prep;14:1—10.

13. Huang Q, Jackson S, Derakhshan S, Lee L, Pham E, Jackson A, et al. Urban-rural differences in COVID-19 exposures and outcomes in the South: a preliminary analysis of South Carolina. PLoS One 2021;16:e0246548.
14. Taylor RA, Haimovich AD. Machine learning in emergency medicine: keys to future success. Acad Emerg Med 2021;28:263—7.
15. Kareemi H, Vaillancourt C, Rosenberg H, Fournier K, Yadav K. Machine learning versus usual care for diagnostic and prognostic prediction in the emergency department: a systematic review. Acad Emerg Med 2021;28:184—96.
16. Raita Y, Goto T, Faridi MK, Brown DFM, Camargo CA, Hasegawa K. Emergency department triage prediction of clinical outcomes using machine learning models. Crit Care 2019;23:64.
17. Magro B, Zuccaro V, Novelli L, Zileri L, Celsa C, Raimondi F, et al. Predicting in-hospital mortality from Coronavirus Disease 2019: a simple validated app for clinical use. PLoS One 2021;16:e0245281.
18. Abdulaal A, Patel A, Al-Hindawi A, Charani E, Alqahtani SA, Davies GW, et al. Clinical utility and functionality of an artificial intelligence—based app to predict mortality in COVID-19: mixed methods analysis. JMIR Form Res 2021;5:e27992.
19. Vittorini P, Casano N, Sinatti G, Santini SJ, Balsano C. The covid-19 decision support system (C19DSS) – a mobile app. In: Proceedings of the 16th international conference on practical applications of computational biology & bioinformatics. (PACBB'22); 2022.
20. World Health Organization. Clinical management of COVID-19; 2022. Available from: https://www.who.int/publications/i/item/WHO-2019-nCoV-clinical-2021-219 [Accessed 1 Feb 2022].
21. Sinatti G, Santini SJ, Tarantino G, Picchi G, Cosimini B, Ranfone F, et al. PaO_2/FiO_2 ratio forecasts COVID-19 patients' outcome regardless of age: a cross-sectional, monocentric study. Intern Emerg Med 2021;17:665—73.
22. Knight SR, Ho A, Pius R, Buchan I, Carson G, Drake TM, et al. Risk stratification of patients admitted to hospital with Covid-19 using the ISARIC WHO Clinical Characterisation Protocol: development and validation of the 4C Mortality Score. BMJ 2020;370:m3339.
23. Haimovich AD, Ravindra NG, Stoytchev S, Young HP, Wilson FP, van Dijk D, et al. Development and validation of the quick COVID-19 severity index: a prognostic tool for early clinical decompensation. Ann Emerg Med 2020;76:442—53.
24. Collins GS, Reitsma JB, Altman DG, Moons K. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement; 2022. Available from: https://core.ac.uk/reader/81583807?utm_source=linkout [Accessed 1 Feb 2022].
25. Little R, Rubin D. Statistical analysis with missing data, 2nd ed. Wiley; 2022. Available from: https://www.wiley.com/en-us/Statistical&tnqx2b;Analysis&tnqx2b;with&tnqx2b;Missing&tnqx2b;Data%2C&tnqx2b;2nd&tnqx2b;Edition-p-9781119013563 [Accessed 12 Mar 2022].
26. Schafer JL. Analysis of incomplete multivariate data. New York: Chapman and Hall/CRC; 1997.
27. Van Buuren S, Groothuis-Oudshoorn K. Mice: multivariate imputation by chained equations in R. J Stat Software 2011;45:1—67.
28. Breiman L, Friedman JH, Olshen RA, Stone CJ. Classification and regression trees. New York: Routledge; 2017.
29. Breiman L. Random forests. Mach Learn 2001;45:5—32.
30. Chen T, Guestrin C. XGBoost: a scalable tree boosting system. In: Proc 22nd ACM SIGKDD int conf knowl discov data min.; 2016: 785—94 pp. Published online [Accessed 13 August 2016].
31. Efron B. Estimating the error rate of a prediction rule: improvement on cross-validation. J Am Stat Assoc 1983;78:316—31.
32. Seiffert C, Khoshgoftaar TM, Van Hulse J, Napolitano A. RUSBoost: improving classification performance when training data is skewed. In: 2008 19th international conference on pattern recognition; 2008:1—4 pp.
33. Gevrey M, Dimopoulos I, Lek S. Review and comparison of methods to study the contribution of variables in artificial neural network models. Ecol Model 2003;160:249—64.
34. Kuhn M. Building predictive models in R using the caret package. J Stat Software 2008;28:1—26.
35. Antonelli M, Penfold RS, Merino J, Sudre CH, Molteni E, Berry S, et al. Risk factors and disease profile of post-vaccination SARS-CoV-2 infection in UK users of the COVID symptom study app: a prospective, community-based, nested, case-control study. Lancet Infect Dis 2022;22:43—55.
36. Bar-On YM, Goldberg Y, Mandel M, Bodenheimer O, Freedman L, Kalkstein N, et al. Protection of BNT162b2 vaccine booster against covid-19 in Israel. N Engl J Med 2021;385:1393—400.
37. Khan M, Mehran MT, Haq ZU, Ullah Z, Naqvi SR, Ihsan M, et al. Applications of artificial intelligence in COVID-19 pandemic: a comprehensive review. Expert Syst Appl 2021;185:115695.
38. Jimenez-Solem E, Petersen TS, Hansen C, Hansen C, Lioma C, Igel C, et al. Developing and validating COVID-19 adverse outcome risk prediction models from a bi-national European cohort of 5594 patients. Sci Rep 2021;11:3246.
39. Noy O, Coster D, Metzger M, Atar I, Shenhar-Tsarfaty S, Berliner S, et al. A machine learning model for predicting deterioration of COVID-19 inpatients. Sci Rep 2022;12:2630.
40. Yadaw AS, Li YC, Bose S, Iyengar R, Bunyavanich S, Pandey G. Clinical features of COVID-19 mortality: development and validation of a clinical prediction model. Lancet Digit Health 2020;2:e516—25.
41. Yao H, Zhang N, Zhang R, Duan M, Xie T, Pan J, et al. Severity detection for the coronavirus disease 2019 (COVID-19) patients using a machine learning model based on the blood and urine tests. Front Cell Dev Biol 2020;8:683.
42. Gianstefani A, Farina G, Salvatore V, Alvau F, Artesiani ML, Bonfatti S, et al. Role of ROX index in the first assessment of COVID-19 patients in the emergency department. Intern Emerg Med 2021;16:1959—65.
43. Vega ML, Dongilli R, Olaizola G, Colaianni N, Sayat MC, Pisani L, et al. COVID-19 Pneumonia and ROX index: time to set a new threshold for patients admitted outside the ICU. Pulmonology 2022;28:13—7.

44. The ARDS Definition Task Force*. Acute respiratory distress syndrome: the Berlin definition. JAMA 2012;307:2526—33.
45. Roca O, Caralt B, Messika J, Samper M, Sztrymf B, Hernandez G, et al. An index combining respiratory rate and oxygenation to predict outcome of nasal high-flow therapy. Am J Respir Crit Care Med 2019;199:1368—76.
46. Ren SY, Wang WB, Gao RD, Zhou AM. Omicron variant (B.1.1.529) of SARS-CoV-2: mutation, infectivity, transmission, and vaccine resistance. World J Clin Cases 2022;10:1—11.

Supplementary Material: This article contains supplementary material (https://doi.org/10.1515/jib-2022-0047).