Predicting the outcome of Coronavirus Disease 2019 (COVID-19) patients: a Machine Learning Systems (MLS) approach

Roberta Garberi[1], Martina Favarato[1,2], Hedwige Gay[1,2], Sara Bragagnolo[1], Fabiana Villa[1], Anna Lucia Zeduri[1], Thomas Langer[1,2], Roberto Fumagalli[1,2]

[1] Department of Medicine and Surgery, University of Milan-Bicocca, Monza, Italy
[2] Department of Anaesthesia and Intensive Care Medicine, Niguarda Ca' Granda, Milan, Italy

BACKGROUND

As of June 19, SARS-CoV-2 has caused in Italy 238,000 cases and 34,500 deaths, the majority being in March and April in Lombardy region.

OBJECTIVE

To test the accuracy of Artificial Intelligence (AI) in predicting the need for ICU (Intensive Care Unit) admission and the outcome of COVID-19 patients using information easily available in every Emergency Department.

METHODS

Patients admitted to the Emergency Department of Niguarda Hospital from February, 23 to April, 17 were screened for eligibility. Inclusion criteria were symptoms compatible with SARS-CoV-2 infection and a positive real time RT-PCR essay. Age < 12 years, absence of evaluation of the leukocyte formula or thoracic imaging, direct dismission from the emergency department and the indication to a limitation of cures constituted exclusion criteria. Analysis was conducted in two phases: first through classic statistics and then through machine learning systems (MLS). For the second part of the analysis we used several algorithms from the Waikato Environment for Knowledge Analysis (WEKA) software applying the stratified 5-fold cross-validation protocol.

Patients admitted to ICU for COVID-19 in Italy and Lombardy. (As of June, 19)

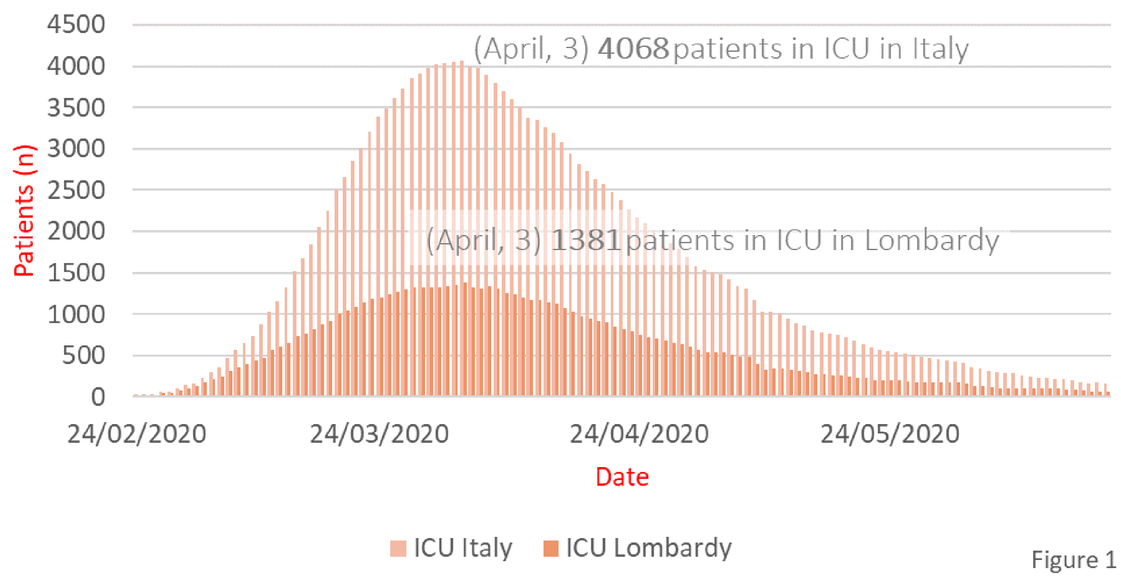

Figure 1

Algorithm	
J48	Classifiers
RandomForest	
NaiveBayes	
MultilayerPerceptron	
Logistic	
AttributeSelectedClassifier	Variable selection
CfsSubsetEval	

Table 1: List of algorithms

RESULTS

323 patients were included in the study. 55 (17%) were admitted to ICU and 37 (11.5%) died. Median age of the overall population was 59 [47-71] years and 62.2% were males. The best machine learning algorithm reached an accuracy of 91.3% in predicting the need for ICU admission and of 88.7% in predicting the outcome.

ICU

Algorithm	Sensitivity (%)	Specificity (%)	Weighted accuracy (%)	AUROC
RandomForest	58	98	91.3	0.870

10 variables selected

5 variables included in both models

- Age
- Blood oxygen saturation
- CRP
- Plasmatic calcium
- Use of medication for mental health

Outcome

Algorithm	Sensitivity (%)	Specificity (%)	Weighted accuracy (%)	AUROC
RandomForest	20	98	88.7	0.840
NaiveBayes	48	89	83.9	0.820

17 variables selected

Tables 2 and 3: ML results (algorithms that showed the best performances) for the two outcomes in study

CONCLUSIONS

This study shows that machine learning systems were able to predict with a good accuracy the two outcomes in study. From these results it may be possible to create a decision helping tool that can support clinicians through the formulation of a risk score of the severe evolution of patients.