Artificial Intelligence helps triage and monitoring of COVID-19 Intensive Care patients

Salah Boussen (salah.boussen@univ-eiffel.fr)
Aix-Marseille University

Pierre-Yves Cordier
Hôpital Laveran

Arthur Malet
Marseille Public University Hospital System

Pierre Simeone
Marseille Public University Hospital System

Sophie Cataldi
Marseille Public University Hospital System

Camille Vaisse
Marseille Public University Hospital System

Xavier Roche
Marseille Public University Hospital System

Alexandre Castelli
Marseille Public University Hospital System

Mehdi Assal
Marseille Public University Hospital System

Guillaume Pépin
Marseille Public University Hospital System

Kevin Cot
Marseille Public University Hospital System

Jean-Baptiste Denis
Marseille Public University Hospital System

Timothée Morales
Marseille Public University Hospital System

Lionel Velly
Aix-Marseille University

Nicolas Bruder
Aix-Marseille University

Research Article

Keywords: artificial intelligence, triage, monitoring, COVID-19, intensive care

Posted Date: April 15th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-307816/v2

Abstract

Background

The rapid spread of coronavirus disease COVID 19 calls for early screening and monitoring of these patients to distinguish those that are likely to worsen from stable patients that may be directed to intermediate care facilities. We designed a score for COVID-19 patients severity assessment, dynamic intubation and prolonged stay prediction using the Breathing Frequency (BF) and oxygen saturation (SPO2) signals.

Methods

We recorded BF, and SPO2 signals of confirmed COVID-19 patients admitted during the first and second outbreak of the pandemic in France (March to May 2020 and September 2020 to February 2021) in an ICU of a teaching hospital. We extracted four features from the signals that represent the four last hours before intubation for intubated patients and the mean of the four hours before the median intubation time for non-intubated patients. These data were used to train AI algorithms for intubation recognition. Algorithm robustness was checked on a validation set of patients. We selected the best algorithm that was applied every hour to predict intubation, thus a severity evaluation. We performed a 24h moving average of these predictions giving a S_{24} severity score that represent the patient's severity during the last 24 h. MS_{24}, the maximum of S_{24} was confronted with the risk of intubation and prolonged ICU stay (>5 days).

Results

We included 177 patients. Among the tested algorithms, the Logistic regression classifier had the best performance. The model had an accuracy of 88.9 % for intubation recognition (AUC=0.92). The accuracy on the validation set was 92.6 %. The S_{24} score of intubated patients was significantly higher than non-intubated patients 48h before intubation and increased 24 hours before intubation. MS_{24} score allows distinguishing three severity situations with an increased risk of intubation: green (3%), orange (30%) and red (76%). A MS_{24} score superior to 20 was highly predictive of an ICU stay greater than 5 day with an accuracy of 88.8% (AUC=0.95).

Conclusions

The score we designed uses simple signals and seems to be efficient to visualize the patient's respiratory situation and may help in decision-making. Real-time computation is easy to implement.

Introduction

The rapid spread of coronavirus disease (COVID 19) challenges Intensive Care Units (ICU) worldwide. The increase in cases may overwhelm ICU capacity[1,2]. This situation calls for early screening and

monitoring of these patients to distinguish patients that are likely to worsen from stable patients that may be directed to intermediate care facilities or regular hospital ward. Here, we present a method that allows a real-time analysis of respiratory signals using an AI algorithm for dynamic severity assessment. Indeed, COVID-19 patients might experience profound and repeated hypoxia without requiring intubation. High flows oxygen therapy proved to be efficient for COVID-19 pneumonia[3–5]. However, some patients may be over treated by non-invasive treatments leading to delayed intubation, hence the interest of early detection of a failure.

Besides intubation prediction, SAPS2 score seems unsuitable for severity assessment of COVID-19 patients. Indeed at ICU entrance, these patients present a single failure that is mild at ICU admission and changes during the first days of the stay. Therefore, rapid and usual scorings fail to predict COVID-19 evolution[6]. An effort to design effective scoring to helps physicians decisions was made recently[7,8]. However, these efficient scores use lab results and may be more challenging to implement. In this study, we present an all automated scoring system that uses only two physiological respiratory signals (Breathing Frequency (BF) and Oxygen Saturation (S_PO_2)) to assess patient's severity, predicting intubation and prolonged length of stay in ICU (> 5 days).

Indeed, early detection of patients that necessitates an extended stay in the ICU may help patients' flow regulation and resource management.

Patients And Methods

In this retrospective, observational study, we included confirmed COVID-19 patients and admitted to our ICU during the first and second outbreak of the pandemic in France (March to May 2020 and September 2020-February 2021). We excluded patients intubated before admission, patients with missing data and patients with a non-intubation decision. The collection of ICU data was approved by the local committee CPP 1 Sud Méditerrannée (RO – 2015/20) and the local computer and freedom committee (CIL). According to french law, this is a non interventional research, therefore participants or family were informed of the research's purpose and their right to decline participation after information by the investigator and delivery of a document containing the information related to the research and the modalities of information of the global results.

1. General Conception of the severity scores

To create a severity score, we used classification algorithms for intubation pattern recognition using simple respiratory signals BF and SPO2. We set the prediction timing at 4 hours before actual intubation. We trained classification algorithms for intubation recognition, and we assumed that intubation was performed for unstable patients that non-invasive therapy was not sufficient anymore. Therefore, the positive prediction of intubation is associated with high severity. This prediction was performed hourly, and we used this prediction that was averaged over 24h to build a score that reflects the patient's actual

severity. Indeed, a patient with a positive prediction of intubation during 24 hours is more severe than one that has one hour or two over 24 hours.

2. Data collection

Physiological data were extracted from the bedside monitors every 5 s. A python script transformed the HL7 format to an exploitable monitor parameters CSV file containing all patient's parameters.

3. Data processing

We used BF and SPO2. These two parameters are recorded for all patients, and may be deployed easily even outside of ICU.

Artefacts as flushing or disconnection were removed using a median filter.

For BF we computed using a sliding window of one hour:

-Maximum Value (BFmax)

-Percentage of Time under 20 breath/minutes (BF20)

For SPO_2 we computed using a sliding window of one hour:

-Minimum Value (SpO_2min)

-Percentage of Time under 90 % (SPO_2-90)

The four parameters vector [BFmax BF20 SpO_2min SPO_2-90] represents patient's state in the last hour.

Features selection was done pragmatically using predictor importance measurement. Features were reduced to the minimum according to parsimony principles. Each hour, the four-parameter state vector described patient's situation.

4. Algorithms training and validation

For non-intubated patients, we computed the mean of the four hours before the median intubation time (80 hours). For intubated patients, we calculated the mean state vector of the four last hours before intubation. Intubation was timed using a capnogramm curve. Indeed, our ICU protocol specifies a mandatory capnogramm for all patient intubation.

Each patient gave a 4 parameters state vector representing the supposed worst condition for intubated patients. We trained the algorithm for severity recognition using the mean state vectors. The number of inputs was 4. Fivefold cross-validation testing determined the consistency of model accuracy.

An independent cohort, including half of the admitted patients, made the final validation. Patients were randomly assigned either to the learning or validation cohort. Learning and test date are provided in

Additional File 1.

5. Scores construction

Patient situation is characterized by its state vector $\mathbf{x}=[a_1,a_2,a_3,a_3]$ that is computed on the last hour (i.e maximum of BF, ratio of time spent under 20 /min of BF, minimum of S_PO_2 and ratio of time spent under 90% of saturation during the last hour). The algorithm is applied to $\mathbf{x}$ giving a prediction $p(\mathbf{x})$

$$p(x) = \begin{cases} 0, \\ 1 \end{cases}$$

This prediction is updated hourly.

The algorithm gives a prediction that the patient will be intubated in four hour, $p(\mathbf{x}) = (1)$ or (0). It reflects the instantaneous state of the patient. However, this state may be due to a movement of the patient (leaving the chair to bed or vice and versa), patient speaking or getting of the high flow nasal canula etc...Therefore we computed the moving average of 24 hours that reflect the patient state on a larger scale.

$$S_{24}(t) = 100.\left(\frac{1}{24}\sum_{i=1}^{i=24} p(t-i)\right)$$

Patients were classified into three categories depending on the maximum MS_{24} of S_{24}.

$$MS_{24}(T) = \max_{24<t<T} S_{24}$$

We excluded the first 24 hours because, at admission, most of the patients are unstable and have a large score that corrects with a support therapy initiation. We also excluded the last 4 hours before intubation because it is the set that was used for algorithm learning. Therefore a patient is in the green category if his score remains in the green category. Orange or red if it S_{24} maximum is in the orange and red category, respectively. However, the patient score may go to a higher severity level but never decrease.

All computations were performed using Matlab® (The Mathworks, Natick, MA, USA). Data, patient example and a script are provided in supplementary files (Additional Files 2 and 3).

Results

1. Population

We included 177 patients (63 [58–71] years, 58 females, 49 deaths). 23 patients were included during the first outbreak and 154 during the second (see flow chart diagrams, Fig. 1). The first 89 patients were used for algorithm training (30 intubations) and 88 as a validation set (29 intubations). Their basic baseline characteristics are listed in Table 1:

Table 1
Patients characteristics

Number	**177**
Age	63 [58–71]
Sexe ratio F/M	58 / 119 (33% / 67%)
Intubation	59 (33 %)
Length of Stay	8 [4–17]
Death	49 (27.6 %)

2. Algorithm accuracy

Among the tested algorithms, linear algorithms have the best performance in the validation set (Table 2).

Table 2
Algorithms performances. AUC: Area under curve, PPV: Positive Predictive Value, Se: Sensibility. Algorithms: LD : Linear Discriminant, LOG Logistic regression classifier, SVM: SupportVector Machine, BN: Bayesian Naïve, CART: Classification And Regression Tree, BAG: bagging tree ensemble, RUS : RUSBosst algorithm, Neural : Neural Network algorithm.

	Training				**Validation**			
	Accuracy	**AUC**	**PPV**	**Se**	**Accuracy**	**AUC**	**PPV**	**Se**
LD	86.5	0.97	85.2	76.7	92.0	0.96	84.4	93.1
LOG	87.6	0.97	85.2	76.7	92.0	0.96	84.4	93.1
SVM	88.8	0.97	83.3	83.3	90.9	0.95	81.8	93.1
BN	88.8	0.96	85.7	80.0	89.8	0.93	81.2	89.7
CART	94.4	0.95	90.3	93.3	87.5	0.90	80.0	82.8
BAG	91.0	0.97	84.0	81.3	89.8	0.93	81.2	89.7
RUS	91.0	0.88	84.0	81.4	84.1	0.91	72.7	82.8
NEURAL	92.1	0.98	89.7	86.7	88.6	0.94	80.6	86.2

We choose the logistic regression classifier due to its performances, generalizability and also because it is easy to deploy and use for the community. Its parameters are listed in Table 3:

Table 3
Logistic regression classifier parameters.

	Estimate	SE	tstat	p-value
Intercept	36,2	15,8	2,2	0,02
BFmax	0,18	0,077	2,4	0,02
BF20	-1,16	2,8	-0,41	0,68
SpO2min	-0,51	0,19	-2,7	0,007
SPO2-90	6,6	3,2	2,05	0,04

3. S_{24} Score

Patients that were intubated had a higher S_{24} score at least 80 hours before intubation (Fig. 2-A), and their score grows 24h before intubation continuously. We found the same difference between intubated and non-intubated patients in the validation cohort, ensuring its validity (Fig. 2-B). The S_{24} score was significantly higher the last 48 hours before intubation and remain high until intubation.

4. MS_{24} score

MS_{24} score is the maximum of S_{24} score, disregarding the first 24 h of the stay. MS_{24} was highly correlated with the occurrence of intubation (Fig. 3-A). The event of intubation grows with MS_{24} score. The validation cohort follows the same law (Fig. 3-A), validating the score. We established a probability law that grows continuously with MS_{24} score (Fig. 3-B). The MS_{24} allows distinguishing three severity situations (green-orange-red) (Table 4). The cut-offs of these categories are established arbitrarily by using the probability law's natural inflexion points corresponding to a cumulative incidence of intubation of 30 and 76%.

Table 4: Number of intubation according to MS_{24} score.

	Score MS_{24}	Number of Patients	Number of Intubations	%
	<25	63	2	3
	25-60	60	18	30
	>60	54	41	76

The MS24 score allows distinguishing between three levels of intubation risk but also between different patients behaviours (Table 4):

Green

($MS_{24} < 25$) corresponds to patients that respond well to the therapeutics, and therefore their hourly score may increase but returns quickly to zero with a low risk of intubation (3 %). Figure 4-A shows a typical S_{24}

evolution of such patients. The patient situation had a mild increase in S_{24} score but returned to a low score. One of the intubated patients was the first admitted patient in our ICU for COVID-19.

Orange

($MS_{24} \leq 60$) corresponds to patients that are unstable with large increases in the hourly score value that return more or less quickly to 0 with an increased risk of intubation (30%)(Fig. 4-A and B). It is difficult to say whether patients intubated in this category were intubated prematurely or, on the contrary, were intubated before they worsened. However, the patient in 4-A stabilized and improved dramatically, while patient in 4-B keeps worsening after a temporary improvement.

Red

($MS_{24} > 60$) corresponds to highly unstable patients with a prolonged higher hourly score and an increased risk of intubation (76 %). Figure 4-C shows the typical evolution of red category patient. The increase of S_{24} score is continuous, and the patient stayed 40 hours at the maximum severity score before intubation.

5. MS_{24} Score and ICU Length of stay

We studied the link between MS_{24} and the length of stay in ICU to use our score for triage. We excluded six intubated patients who died prematurely due to uncontrolled situation (cardiac arrest and pulmonary embolisms) for this study. MS_{24} was correlated to the ICU length of stay (Fig. 5-A). We found that a MS_{24} score superior to 20 was highly predictive of an ICU stay greater than 5 day with an accuracy of 88.8% (PPV = 93.0%, Sensibility = 90.7%, AUC = 0.95) (Fig. 5B and C). For intubated patients, 58/59 patients had a S_{24} score greater than 20 in the first 24 hours. Therefore, a prolonged ICU was predicted on the first day.

Discussion

We used BF and SPO2 to assess COVID-19 patients severity by mean of the IA algorithms. The algorithm output was used to design two scores S_{24} and MS_{24}. S_{24} is computed continuously and reflects the patient's state, while MS_{24} reflects the patient's absolute severity and helps to categorize the patient. The latest may allow for patients monitoring and decision-making by dividing patients into three categories green, orange and red. First, green patients have a lower risk of intubation. They may be directed to intermediate care, while others should stay monitored carefully because of a high intubation risk. Secondly, COVID-19 patients are highly hypoxic and experience during their stay profound desaturation episodes that may lead to premature intubation in a constrained pandemic situation[9] all the more that intubation is at risk and debated[10,11]. During an episode of desaturation, a low S_{24} may help to postpone intubation. Finally, intubation should be considered for a red patient with a continuously increasing score. Indeed, prompt detection of non-invasive treatment failure is of great importance to not delay intubation. Recent works suggest that patients' self-inflected injury during spontaneous breathing is

a possible mechanism of aggravating lung damage and is responsible for increased mortality[12]. However, this question is under debate[13,14].

The three categories distinguish three patients behaviour under non-invasive therapy. Indeed, green patients characterize patients that respond well. Their instantaneous score (i.e. algorithm classification) sometimes rise to 1 during effort or nursing, but it returns quickly to zero. In the orange category, the algorithm prediction may stay at 1 for a longer time, reflecting a high instability that may decompensate leading to intubation. Finally, the red category characterizes highly unstable patients that are not stabilized by non-invasive therapy.

The utility of such scores is multiple. Indeed, it helps resource management in a constraint period: Green patient may be taken care in intermediate facilities or even in regular floor with adequate non-invasive oxygen therapy and simple SpO_2 and BF monitoring. The same algorithm may be used then with smaller devices used routinely to continuously monitor regular floor patients[15,16]. Secondly, early detection of a prolonged stay in ICU helps resource management. Indeed, a high number of bed blocking patients may trigger the opening of new premises [17] or evacuation of patients of remote regions[18].

A second utility of this score may be to detect earlier the most severe patients at ICU entrance and may be used for initiating future therapies for those patients.

Finally, this score may be used to compare patient's severity as SpO_2 and BF are recorded routinely.

This pilot study has limitations. First, we focalized the algorithm on COVID-19 patients, and we did not consider all patients at risk of intubation. A universal intubation prediction algorithm could have been developed and then calibrated for COVID-19 patients. However, it seems to us that our policy for intubation of covid-19 patients is different, especially since our centre is mixed intensive care with very heterogeneous patients.

Secondly, this monocentric study reflects our intubation policy of COVID-19 patients. However, we chose the Logistic Regression Classifier because it is accurate but also very easy to modify and adapt for other centres. We provide our data set (Additional File 1). For modification, one could add its own data and then modify the coefficient of probability computation. We also provide a script and an example signal to compute S_{24} and MS_{24} (Additional Files 2 and 3). We can also update continuously the algorithm by training it on all previous patients.

Conclusion

The score we designed uses simple signals and seems to be efficient to visualize the patient's respiratory situation and may help in decision-making. Real-time computation is easy to implement, allowing forecasting probability of tracheal intubation and prolonged stay in ICU for COVID-19 patients. The algorithm may be also used in regular floors when patients are monitored with continuous portable

devices. The use of the algorithm for non-COVID- 19 patients who may worsen their respiratory situation needs a dedicated study.

References

[1] Maves RC, Downar J, Dichter JR, Hick JL, Devereaux A, Geiling JA, 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;158:212–25. https://doi.org/10.1016/j.chest.2020.03.063.

[2] Tan E, Song J, Deane AM, Plummer MP. Global Impact of Coronavirus Disease 2019 Infection Requiring Admission to the ICU: A Systematic Review and Meta-analysis. CHEST 2020;0. https://doi.org/10.1016/j.chest.2020.10.014.

[3] Xia J, Zhang Y, Ni L, Chen L, Zhou C, Gao C, et al. High-Flow Nasal Oxygen in Coronavirus Disease 2019 Patients With Acute Hypoxemic Respiratory Failure: A Multicenter, Retrospective Cohort Study*. Critical Care Medicine 2020;48:e1079. https://doi.org/10.1097/CCM.0000000000004558.

[4] Demoule A, Vieillard Baron A, Darmon M, Beurton A, Géri G, Voiriot G, et al. High-Flow Nasal Cannula in Critically Ill Patients with Severe COVID-19. Am J Respir Crit Care Med 2020;202:1039–42. https://doi.org/10.1164/rccm.202005-2007LE.

[5] Calligaro GL, Lalla U, Audley G, Gina P, Miller MG, Mendelson M, et al. The utility of high-flow nasal oxygen for severe COVID-19 pneumonia in a resource-constrained setting: A multi-centre prospective observational study. EClinicalMedicine 2020;28. https://doi.org/10.1016/j.eclinm.2020.100570.

[6] Hu H, Yao N, Qiu Y. Comparing Rapid Scoring Systems in Mortality Prediction of Critically Ill Patients With Novel Coronavirus Disease. Acad Emerg Med 2020;27:461–8. https://doi.org/10.1111/acem.13992.

[7] Liang W, Liang H, Ou L, Chen B, Chen A, Li C, et al. Development and Validation of a Clinical Risk Score to Predict the Occurrence of Critical Illness in Hospitalized Patients With COVID-19. JAMA Intern Med 2020;180:1081–9. https://doi.org/10.1001/jamainternmed.2020.2033.

[8] Burdick H, Lam C, Mataraso S, Siefkas A, Braden G, Dellinger RP, et al. Prediction of respiratory decompensation in Covid-19 patients using machine learning: The READY trial. Comput Biol Med 2020;124:103949. https://doi.org/10.1016/j.compbiomed.2020.103949.

[9] Dar M, Swamy L, Gavin D, Theodore A. Mechanical Ventilation Supply and Options for the COVID-19 Pandemic: Leveraging All Available Resources for a Limited Resource in a Crisis. Annals ATS 2020. https://doi.org/10.1513/AnnalsATS.202004-317CME.

[10] Hyman JB, Leibner ES, Tandon P, Egorova NN, Bassily-Marcus A, Kohli-Seth R, et al. Timing of Intubation and In-Hospital Mortality in Patients With Coronavirus Disease 2019. Critical Care Explorations

2020;2:e0254. https://doi.org/10.1097/CCE.0000000000000254.

[11] Matta A, Chaudhary S, Bryan Lo K, DeJoy RI, Gul F, Torres R, et al. Timing of Intubation and Its Implications on Outcomes in Critically Ill Patients With Coronavirus Disease 2019 Infection. Critical Care Explorations 2020;2:e0262. https://doi.org/10.1097/CCE.0000000000000262.

[12] Grieco DL, Menga LS, Eleuteri D, Antonelli M. Patient self-inflicted lung injury: implications for acute hypoxemic respiratory failure and ARDS patients on non-invasive support. Minerva Anestesiol 2019;85:1014–23. https://doi.org/10.23736/S0375-9393.19.13418-9.

[13] Tobin MJ, Laghi F, Jubran A. P-SILI is not justification for intubation of COVID-19 patients. Ann Intensive Care 2020;10:105. https://doi.org/10.1186/s13613-020-00724-1.

[14] Gattinoni L, Marini JJ, Busana M, Chiumello D, Camporota L. Spontaneous breathing, transpulmonary pressure and mathematical trickery. Ann Intensive Care 2020;10. https://doi.org/10.1186/s13613-020-00708-1.

[15] Weenk M, Koeneman M, van de Belt TH, Engelen LJLPG, van Goor H, Bredie SJH. Wireless and continuous monitoring of vital signs in patients at the general ward. Resuscitation 2019;136:47–53. https://doi.org/10.1016/j.resuscitation.2019.01.017.

[16] Downey C, Ng S, Jayne D, Wong D. Reliability of a wearable wireless patch for continuous remote monitoring of vital signs in patients recovering from major surgery: a clinical validation study from the TRaCINg trial. BMJ Open 2019;9. https://doi.org/10.1136/bmjopen-2019-031150.

[17] Verdonk F, Garnier M, Bonnet F, Jabaudon M. The “ephemeral” intensive care units that saved the French healthcare system, a new concept to be preserved. Eur J Anaesthesiol 2021;38:441–2. https://doi.org/10.1097/EJA.0000000000001320.

[18] Nguyen C, Montcriol A, Janvier F, Cungi P-J, Esnault P, Mathais Q, et al. Critical COVID-19 patient evacuation on an amphibious assault ship: feasibility and safety. A case series. BMJ Mil Health 2020. https://doi.org/10.1136/bmjmilitary-2020-001520.

Figures

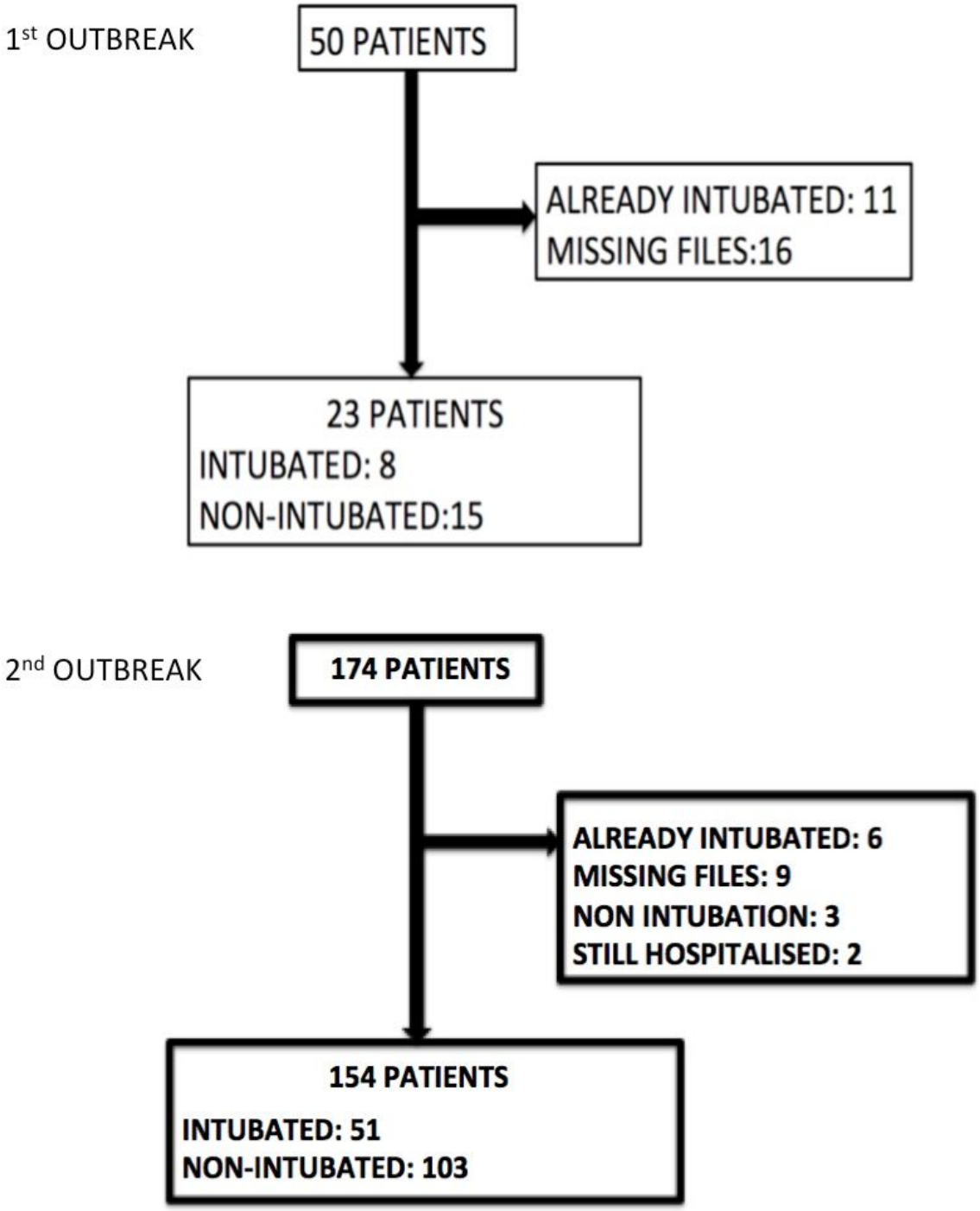

Figure 1

Study Flow chart. Patients were included during the first outbreak (March to May 2020, and the second September 2020 to February 2021).

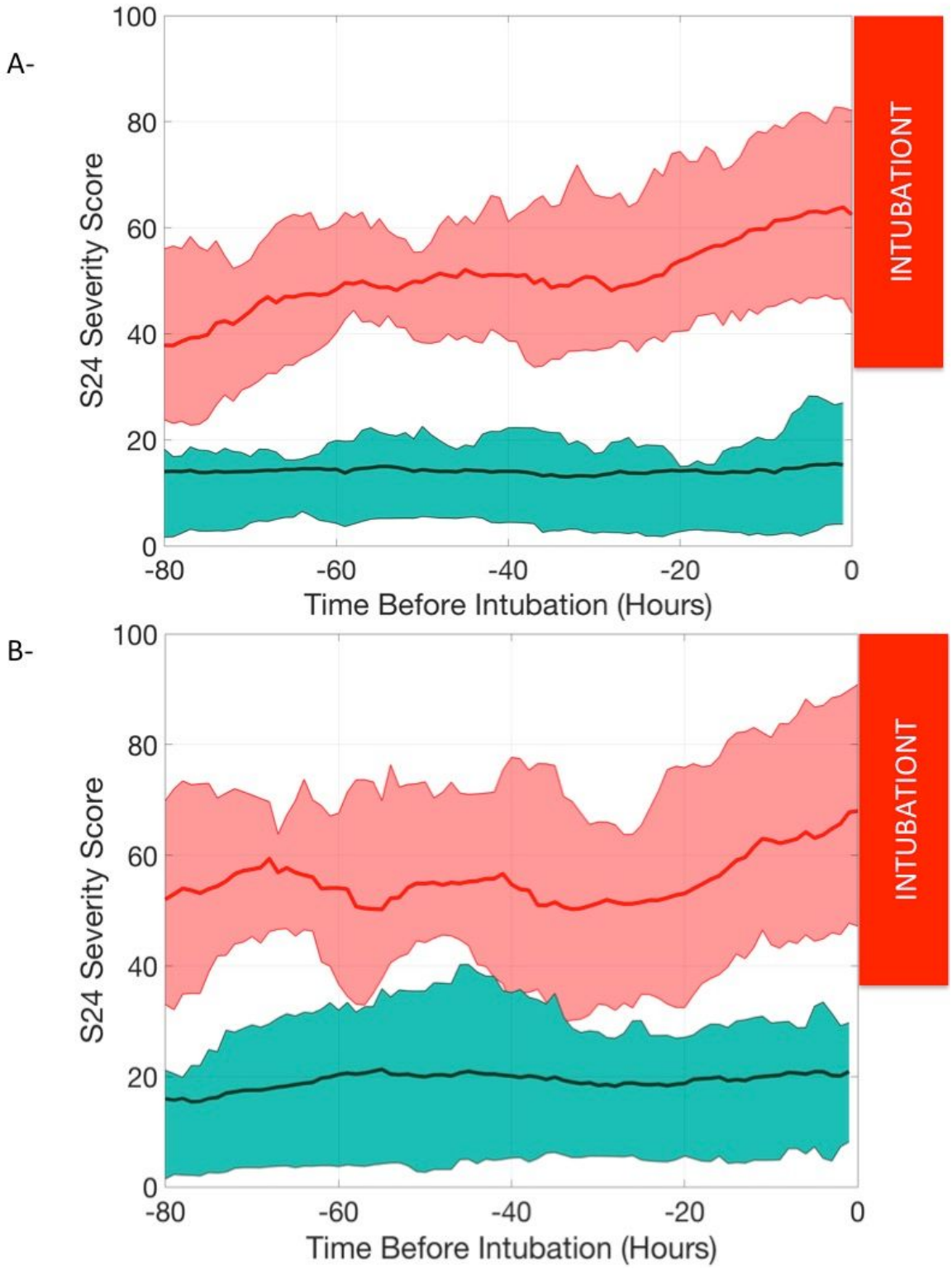

Figure 2

A- Training (A) and Validation (B) cohort dynamic changes of S24 score 48h before intubation. patients from the intubated group are in red. The Plain thick line is mean and fine lines are 25%-75% interval. The green area represents the non-intubated group, the dashed thick line is mean and fine lines correspond to the 25-75% confidence interval. S24 score discriminates both group at least 80 h before intubation. We note a net increase of S24 score 24 h before intubation.

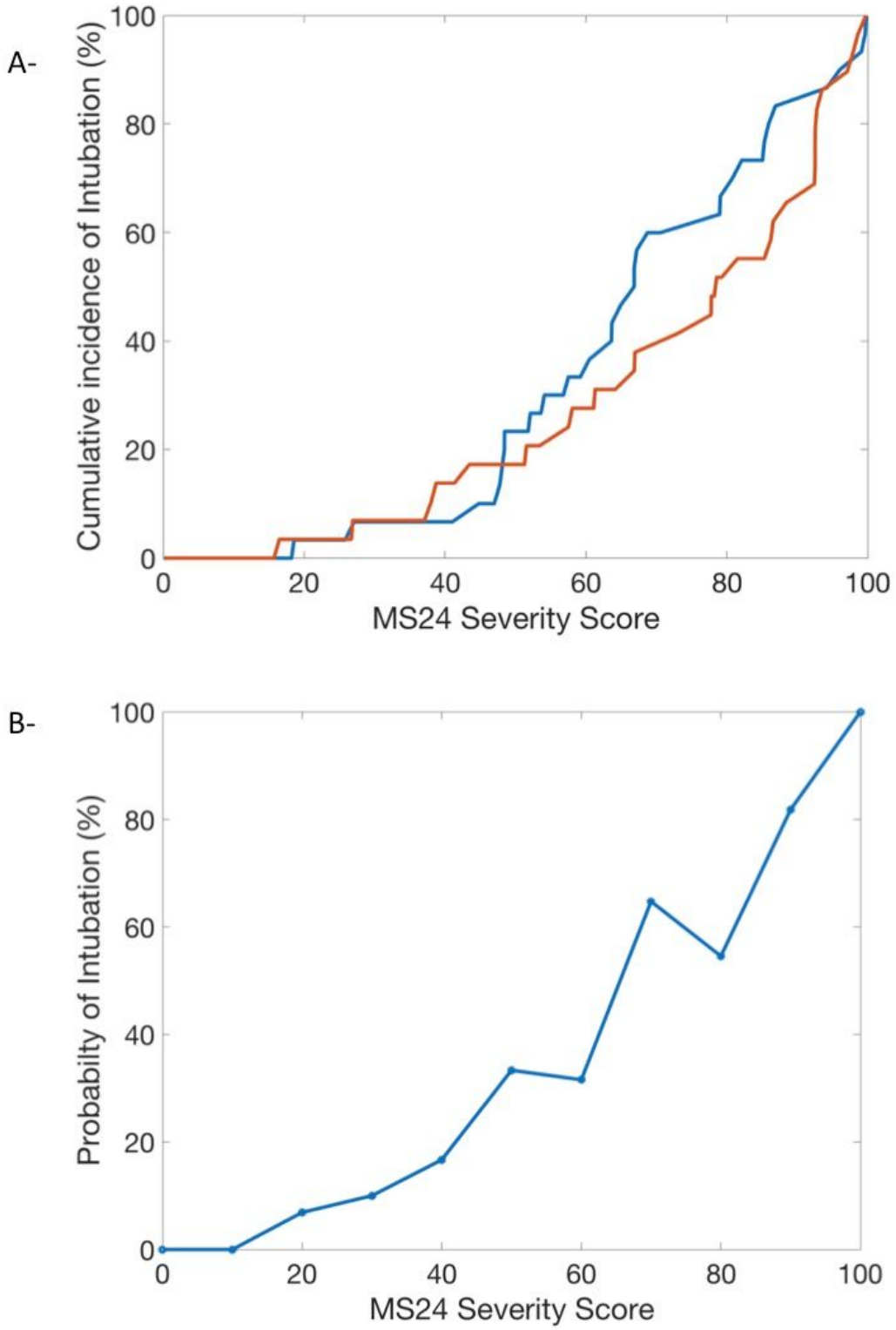

Figure 3

A-Evolution of the cumulative incidence of intubation with the severity score was expressed as the percentage of intubation. The blue line is the training cohort and the red line, the validation cohort. B-Evolution of the probability of intubation according to the MS24 score. The probability grows continuously until reaching 1 for MS24=100.

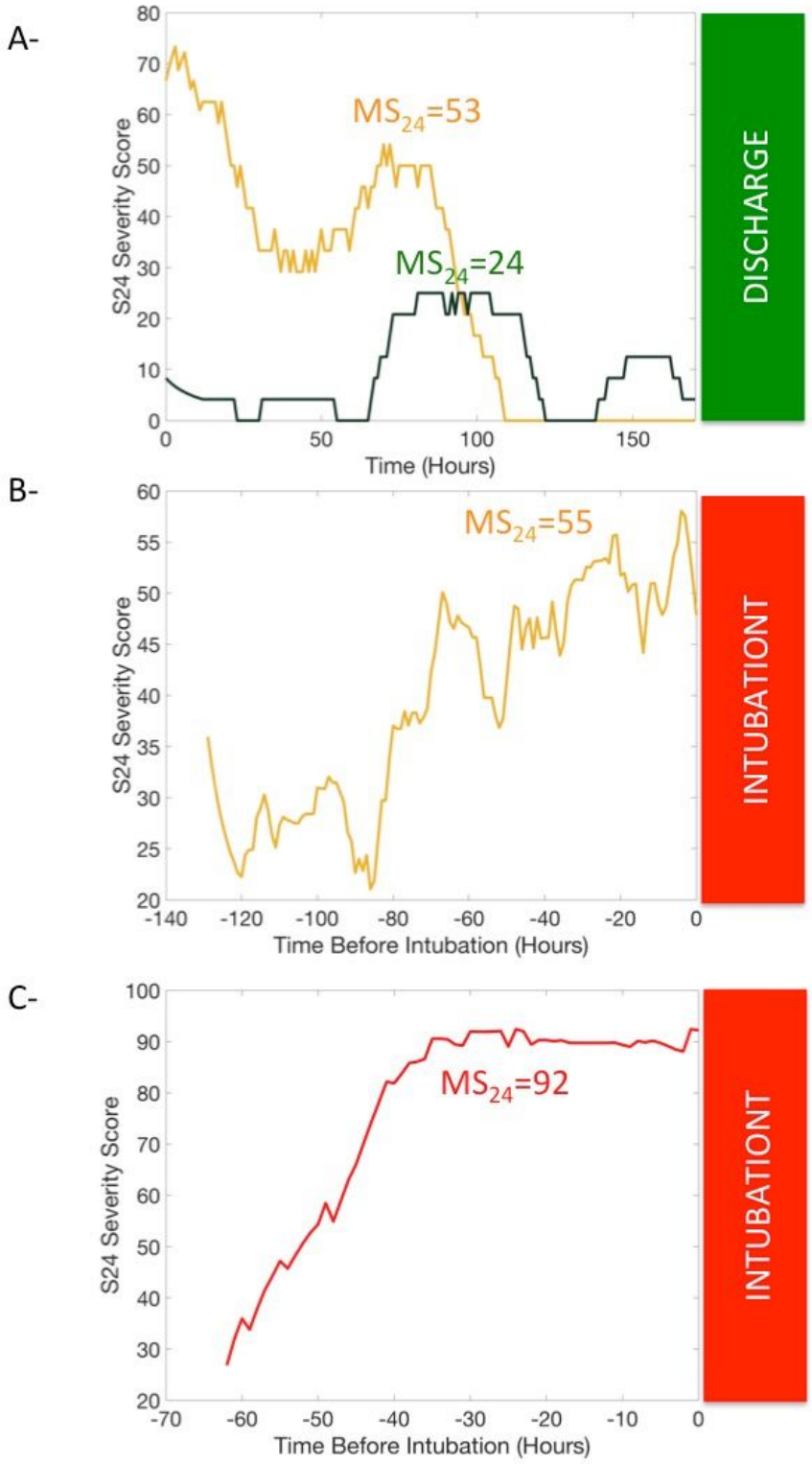

Figure 4

A- Score evolution of a stable patient (green). The score S24 may rise but return quickly to zero, staying under 25. The orange line shows an unstable patient with large score increases. This patient condition had improved without tracheal intubation. B : Unstable patient with a prolonged stay in the warning orange zone. This patient had to be intubated C: Unstable patient with a terminal increase to score 90, 40 h before intubation.

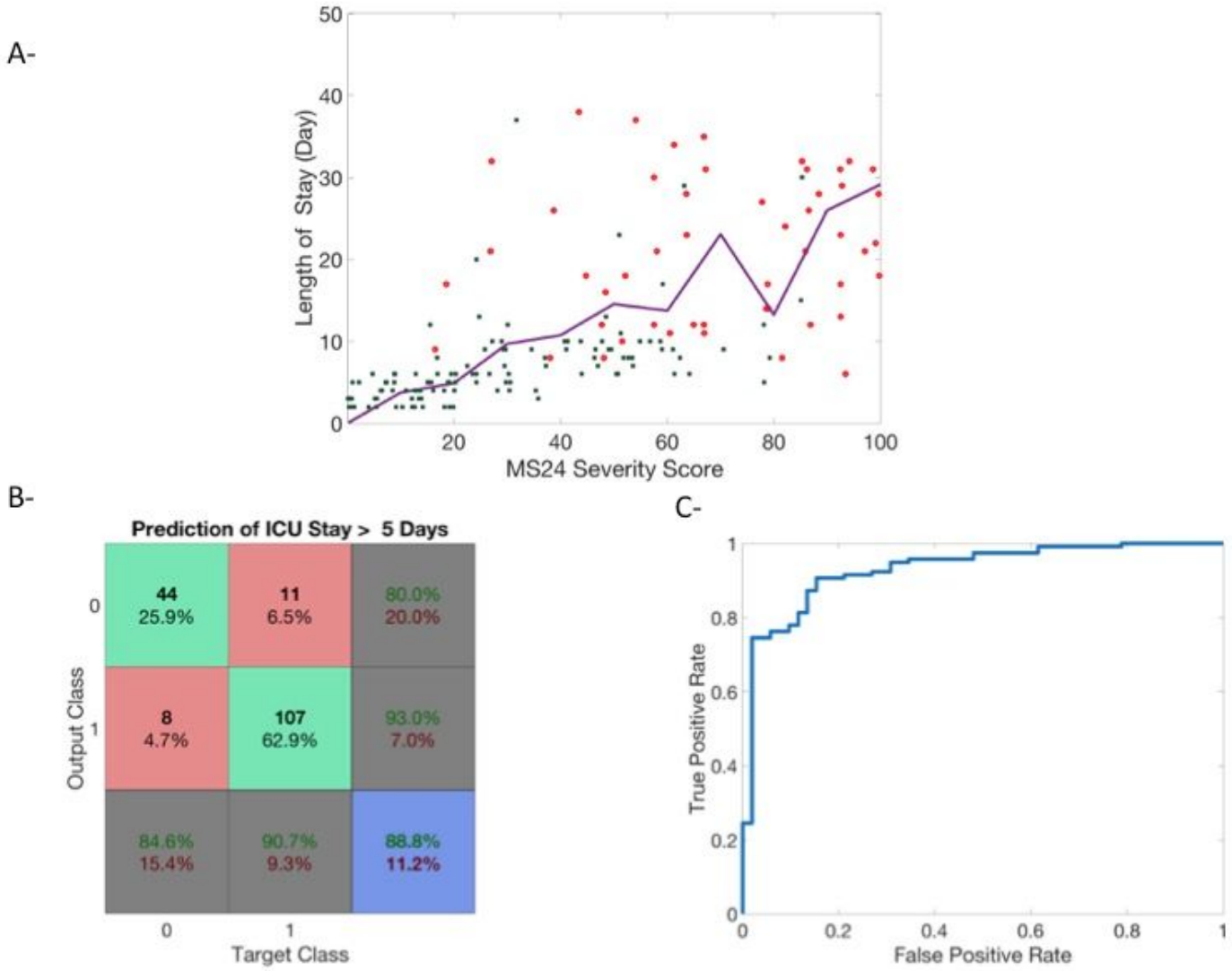

Figure 5

A-Changes of Length of Stay with MS24 score. (Lengths of stay greater than50 days were not plotted). Green dots are non-intubated patients, and red dots are intubated patients. Purple line is the moving average of the length of stay. B-Confusion matrix of a predication of length of stay > 5 days using a MS24 score greater than 20 that corresponds to the optimum of the ROC curve plotted in figure C.

Supplementary Files

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

- DATA.xlsx
- example.xlsx
- ScoreComputation.m
- floatimage1.jpeg