A novel high specificity COVID-19 screening method based on simple blood exams and artificial intelligence

Authors:

1 - Felipe Soares, M.Phil (Corresponding Author) fs (at) felipesoares (dot) net

- University of Sheffield (Department of Computer Science) - United Kingdom
- Universidade Federal do Rio Grande do Sul (Department of Industrial Engineering) - Brazil

2 - Aline Villavicencio, Ph.D

- University of Sheffield (Department of Computer Science) - United Kingdom
- Universidade Federal do Rio Grande do Sul (Institute of Informatics) - Brazil

3 - Michel José Anzanello, Ph.D
- Universidade Federal do Rio Grande do Sul (Department of Industrial Engineering) - Brazil

4 - Flávio Sanson Fogliatto, Ph.D
- Universidade Federal do Rio Grande do Sul (Department of Industrial Engineering) - Brazil

5 - Marco A. P. Idiart, Ph.D
- Universidade Federal do Rio Grande do Sul (Department of Physics) - Brazil

6 - Mark Stevenson, Ph.D
- University of Sheffield (Department of Computer Science) - United Kingdom

Article will be submitted to The Lancet series

A novel high specificity COVID-19 screening method based on simple blood exams and artificial intelligence

Authors: Felipe Soares, Aline Villavicencio, Michel José Anzanello, Flávio Sanson Fogliatto, Marco Idiart, Mark Stevenson

Summary

Background

The SARS-CoV-2 virus responsible for COVID-19 poses a significant challenge to healthcare systems worldwide. Despite governmental initiatives aimed at containing the spread of the disease, several countries are experiencing unmanageable increases in the demand for ICU beds, medical equipment, and larger testing capacity. Efficient COVID-19 diagnosis enables healthcare systems to provide better care for patients while protecting caregivers from the disease. However, many countries are constrained by the limited amount of test kits available, the lack of equipment and trained professionals. In the case of patients visiting emergency rooms (ERs) with a suspect of COVID-19, a prompt diagnosis can improve the outcome and even provide information for efficient hospital management. In this context, a quick, inexpensive and readily available test to perform an initial triage at ER could help to smooth patient flow, provide better patient care, and reduce the backlog of exams.

Methods

In this Case-control quantitative study, we developed a strategy backed by artificial intelligence to perform an initial screening of suspect COVID-19 cases. We developed a machine learning classifier that takes widely available simple blood exams as input and predicts if that suspect case is likely to be positive (having SARS-CoV-2) or negative(not having SARS-CoV-2). Based on this initial classification, positive cases can be referred for further highly sensitive testing (e.g. CT scan, or specific antibodies).

We used publicly available data from the Albert Einstein Hospital in Brazil from 5,644 patients. Focussing on using simple blood exams, a sample of 599 subjects that had the fewest missing values for 16 common exams were selected. From these 599 patients, only 81 were positive for SARS-CoV-2 (determined by RT-PCR).

Based on this data, we built an artificial intelligence classification framework, ER-CoV, aiming at determining which patients were more likely to be negative for SARS-CoV-2 when visiting an ER and that were categorized as a suspect case by medical professionals.

The primary goal of this investigation is to develop a classifier with high specificity and high negative predictive values, with reasonable sensitivity.

Findings

We identified that our framework achieved an average specificity of 92.16% [95% CI 91.73 - 92.59] and negative predictive value (NPV) of 95.29% [95% CI 94.65% - 95.90%]. Those values are completely aligned with our goal of providing an effective low-cost system to triage suspected patients at ERs. As for sensitivity, our model achieved an average of 63.98% [95% CI 59.82% - 67.50%] and positive predictive value (PPV) of 48.00% [95% CI 44.88% - 51.56%].

An error analysis identified that, on average, 45% of the false negative results would have been hospitalized anyway, thus the model is making mistakes for severe cases that would not be overlooked, partially mitigating the fact that the test is not high-sensitive.

All code for our AI model, called ER-CoV is publicly available at https://github.com/soares-f/ER-CoV.

Interpretation

Based on the capacity of our model to accurately predict which cases are negative from suspected patients arriving at emergency rooms, we envision that this framework can play an important role in patient triage. Probably the most important outcome is related to testing availability, which at this point is extremely low in many countries. Considering the achieved specificity, we would reduce by at least 90% the number of SARS-CoV-2 tests performed at emergency rooms, with the chance of getting a false negative at around 5%. The second important outcome is related to patient management in hospitals. Patients predicted as positive by our framework could be immediately separated from the other patients while waiting for the results of confirmatory tests. This could reduce the spread rate inside hospitals since in many hospitals all suspected cases are kept in the same ward. In Brazil, where the data was collected, rate infection is starting to quickly spread, the lead time of a SARS-CoV-2 can be up to 2 weeks.

Funding

University of Sheffield provided financial support for the Ph.D scholarship for Felipe Soares

Introduction

The SARS-CoV-2 virus responsible for Covid19 has posed a significant challenge to healthcare systems worldwide (Phelan et al., 2020). Despite governmental initiatives aimed at containing the spread of the disease, several countries are experiencing unmanageable increases in the demand for ICU beds, medical equipment, and larger testing capacity. By April 5, 2020, more than 1.2 million people were infected by the new coronavirus, with over 60,000 deaths according to the World Health Organization (WHO) situation report 76[1].

Efficient COVID-19 diagnosis enables healthcare systems to provide better care to patients while protecting caretakers from the disease. Most tests for the SARS-CoV-2 virus responsible for COVID-19 may either (i) detect the presence of a virus or a protein, called molecular test, or (ii) detect antibodies produced as a reaction for virus exposure (Li et al., 2020). Tests of type (i) are usually related to Polymerase Chain Reaction (PCR) which is labor-intensive in terms of laboratory procedures (Corman et al., 2020). Tests of type (ii) usually detect the IgG and IgM immunoglobulins related to SARS-CoV-2 and can be commercialized in the form of rapid tests (Li et al., 2020). However, the global clinical industry is currently incapable of meeting such demand. For instance, in Brazil, with a population of more than 200 million people, according to the Health Ministry[2], the country received at the end of March/2020 only 500,000 rapid tests.

The frontline of medical care for COVID-19 is the emergency rooms at hospitals or health centers, which have to identify patients with COVID-19 from the ones with similar symptoms but that present other respiratory diseases, such as fever, cough, dyspnoea, and fatigue (Lake, 2020). The quick determination of patient status regarding COVID-19 may determine follow-up procedures that can improve overall patient outcome, as well as protect medical professionals. Thus, a quick, inexpensive and broadly available testing for this scenario is of utmost interest.

Research in context

Evidence before this study

We searched PubMed on April 7, 2020, for studies using the terms ("SARS-CoV-2" OR "COVID-19" OR “coronavirus”) AND ("artificial intelligence" or "machine learning" or "data science") without any restrictions regarding language or article type. We found no articles describing predictive methods (tests) for suspected patients based only on blood components. Articles regarding artificial intelligence in this context were found for automatic

---

1 https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200405-sitrep-76-covid-19.pdf?sfvrsn=6ecf0977_2

2 https://www.saude.gov.br/noticias/agencia-saude/46623-brasil-inicia-a-distribuicao-de-500-mil-testes-rapidos

assessment of chest CT. Previous studies correlated the C reactive protein to viral infections, mildly elevated alanine aminotransferase (ALT) in acute respiratory distress syndrome (ARDS) patients, leukopenia and lymphopenia.

Added value of this study

This is the first study to report a predictive system with high specificity and acceptable sensibility for the triage of suspected cases of COVID-19 in emergency rooms using only simple blood exams. We achieved an average specificity of 92.16% with negative predictive value of 95.29%. The average rate of false negatives is only 4.06%.

More specific tests are already available in the market, but at the current moment of the SARS-CoV-2 pandemic, stocks are short and result delivery time is long. With our solution, practically any clinical laboratory would be able to produce the information that is used as input in our method (ER-CoV) at very low cost.

Implications of all the available evidence

Our findings support that it would be possible to reduce at least 90% the number of SARS-CoV-2 tests performed at emergency rooms, with the chance of getting a false negative at around 5%, thus making screening more accessible for the population. Secondary implications are related to patient management, since with daily blood samples it would be possible to track patients that may have developed COVID-19 inside the hospital, then allowing better facility isolation.

Artificial Intelligence (AI) methods have already been used in other medical scenarios, such as for the detection of colorectal cancer using blood plasma (Soares et al., 2017), prediction of drug-plasma binding (Kumar et al., 2018), and identification of patients with atrial fibrillation during sinus rhythm (Attia et al., 2019). In the field of metabolomics, AI also plays an important role (Bahado-Singh et al., 2019).

Considering the aforementioned successes in integrating AI and medicine, we propose ER-CoV, an artificial intelligence-based screening method that uses blood exams to triage patients suspect of COVID-19 arriving at emergency rooms.

Methods

Study population

We performed a retrospective case-control study on data collected by the Brazilian Albert Einstein Hospital, a leading facility in the management of COVID-19 in the state of São Paulo. The data is of public access[3] and anonymized, encompassing 5,644 patients, and 108 exams.

According to the data provider, the anonymized data was collected from patients seen at the hospital's emergency room who were initially suspected cases of COVID-19 according to the hospital's workflow for SARS-CoV2[4]. The main criteria by the latest guidelines (01/04/2020) are SpO2 >= 93% or altered pulmonary auscultation; and lack of clinical deterioration or organ dysfunction.

From the 108 variables available in the dataset, we devoted our research to the ones that can be quickly obtained in most laboratories:

1. Mean platelet volume
2. Leukocytes
3. Mean corpuscular volume (MCV)
4. Creatinine
5. Red blood cells
6. Basophils
7. Monocytes
8. Potassium
9. Lymphocytes
10. Mean corpuscular hemoglobin concentration (MCHC)
11. Red blood cell distribution width (RDW)
12. Sodium
13. Mean corpuscular hemoglobin concentration (MCHC)
14. Eosinophils
15. C reactive protein
16. Urea

The reference method for determining if the individual was positive for SARS-CoV-2 was reported to be the reverse-transcriptase polymerase chain reaction (RT-PCR) method as described in the national Brazilian guidelines[5].

---

[3] https://www.kaggle.com/einsteindata4u/covid19

[4] https://medicalsuite.einstein.br/pratica-medica/Documentos%20Doencas%20Epidemicas/Manejo-de-casos-suspeitos-de-sindrome-respiratoria-pelo-COVID-19.pdf - In Portuguese

[5] https://portalarquivos2.saude.gov.br/images/pdf/2020/marco/04/2020-03-02-Boletim-Epidemiol--gico-04-corrigido.pdf

Proposed workflow for patient screening using artificial intelligence

The scenario we developed our solution for is for patients arriving at a hospital's emergency room. In Figure 1 we depict our proposed workflow for the usage of this artificial intelligence (AI) model.

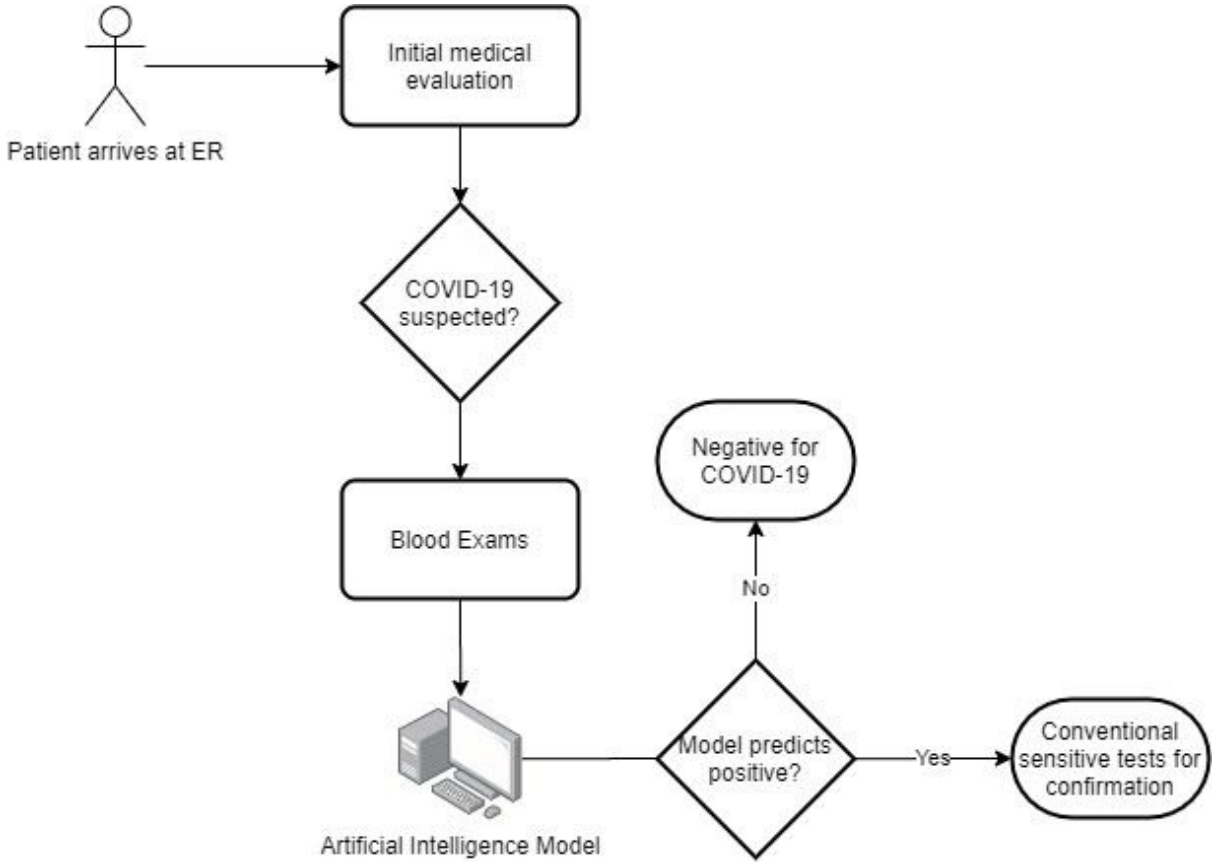

Figure 1: Diagram of our envisioned workflow using the ER-CoV artificial intelligence model in a hospital emergency room (ER).

A patient would initially be evaluated by health professionals to determine if she/he is a probable case of COVID-19.

Once the patient is determined as a suspect case, simple blood exams would be requested, since they are the input for the ER-CoV artificial intelligence model. Our model uses the blood exams mentioned in the study population section.

The results of these exams are used as input to a pre-trained ER-CoV model that will output a positive or negative result. In the case of a positive result, due to its relatively lower sensitivity, we recommend that a second test be performed, such as PCR based, antibodies

rapid kits, or CT scan. If the result is negative, the patient is very likely to not have COVID-19.

In a global scenario of an extreme shortage of testing kits and of laboratory supplies for PCR related exams, we foresee that our proposed method can have a great impact on patient diagnosis at ER, especially in developing and low income countries.

Proposed artificial intelligence method for COVID-19 triage at ERs (ER-CoV)

The AI method requires a series of blood exams already listed and uses advanced techniques to model healthy subjects and predict the status of new patients regarding COVID-19. The overall flowchart of our method (ER-CoV) is shown in Figure 2.

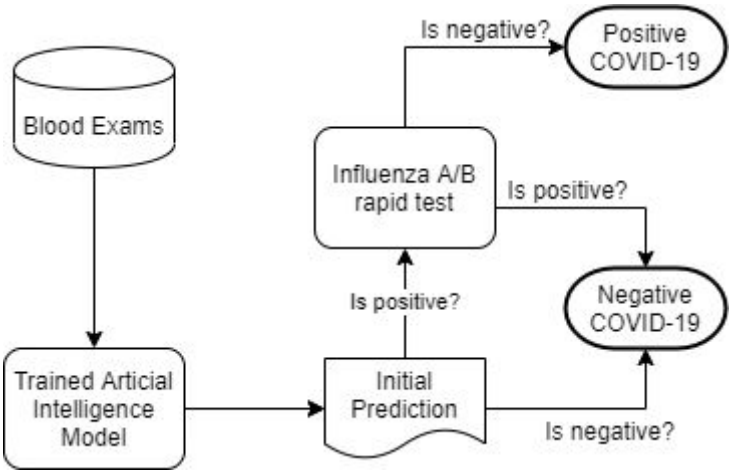

Figure 2: Flowchart for ER-CoV, our AI based method for prediction of COVID-19 in suspected patients in ERs. The output from the AI model may not be final if its output is positive.

In this case, the blood exams are fed to ER-CoV, which will provide an initial prediction of COVID-19 status for the suspect ER patient. As found in the dataset, this initial prediction may be misclassified with other respiratory infections, especially influenza. To reduce possible false positives, we included an intermediary step if a “positive” result is predicted: rapid testing for Influenza A and Influenza B since they are common respiratory diseases with some of the same symptoms. If the result is positive for any of the two Influenza strains, the final prediction will then be reclassified as “negative”. In the dataset there were no positive cases of both Influenza A or B and COVID-19.

If Influenza tests are not available, one could consider the case as “positive”. Analogously, if other additional rapid tests are available, such as for H1N1

Rhinovirus/Enterovirus, they could be added to the intermediary step for differential diagnosis before deeming the sample as positive.

The AI model is trained using a combination of three techniques: Support Vector Machines (Boser et al., 1992), SMOTEBoost (Chawla et al., 2003), and ensembling (Rokach, 2010). Before training our model, C reactive protein missing values (99 out of 599) were imputed using the kNN algorithm (Torgo, 2016) (k=5).

Support Vector Machines (SVM) are an AI technique that can be used for classification (i.e. given a sample, predict if that sample is “positive” or “negative”). This algorithm has been successfully used in many biomedical applications and tends to present good performance. However, due to the small number of “positive” samples in the dataset (13.52%), the algorithm is prone to favor the majority class, that is, predicting all samples as “negative”, resulting in extremely low sensitivity. Oversampling and ensemble methods were applied to mitigate against this problem. SMOTE (Synthetic Minority Oversampling Technique) takes account of the relatively small number of positive examples by creating additional synthetic examples of this class.

The boosting algorithm identifies the incorrect predictions and assigns a larger weight for that sample, such that at the next iteration the algorithm will pay more attention to it. SMOTEBoost combines both approaches to improve classification performance for the smaller class.

Ensembling by its turn is based on the assumption that when one combines a collection of predictions from different models, the final prediction will have better performance. In our case, we trained 200 SVM-based SMOTEBoost models. The initial prediction of our AI model is the average probability for all the 200 models, which is given as “positive” if the probability is greater than 0.5, otherwise, the model will predict the sample as “negative”. Details about model training are provided in the supplementary material, as is the source code.

Incremental developments will be updated at https://github.com/soares-f/ER-CoV.

Numerical Analysis

For the statistical analysis, we computed the following statistics: Sensitivity, Specificity, Positive Predictive Value (PPV), Negative Predictive Value (NPV) (Kirkwood & Sterne, 2010).

AI models may be affected by a specific partition of the data, that is, the particular samples used to calibrate and evaluate the model (the training and test sets). To counter that, we repeated the process of training and evaluation 100 times using different partitions of data for training and testing and storing all information from each run. Approximately 90% of the data was assigned for training, and 10% for testing. Partitioning of the data training and test was carried out using stratified random sampling r to ensure that training and test samples have approximately the same proportion of positives and negatives.

After running these steps for ER-CoV, as depicted in Figure 2, we computed the average of the aforementioned statistics and their 95% confidence interval evaluated through BCa bootstrapping (Efron, 1987) with 500 replicates. All models and statistics were carried out in R version 3.6.3.

Results

Characterization of patients

A total of 599 patients were included in this study, corresponding to a subsample of the original dataset of 5,644 individuals. The subsampling was performed to select only patients that have reported blood exams.

Due to the anonymization process performed by the hospital, we do not have access to the nominal values of the exams, but their standardized values (i.e. zero mean and unit standard deviation). The same is true for patient age, which was split in quantiles but no additional information was given regarding this data, therefore we did not rely on age in any of our computations. Gender, race, and possible time of infection were not available.

A total number of 81 positive cases were present in the dataset, while 518 were negatives, giving a prevalence of 13.52%. When computing the prevalence in the original dataset, with the missing values, we found a prevalence of 9.88%. In Figure 3, we present the diagram showing the flow of patients.

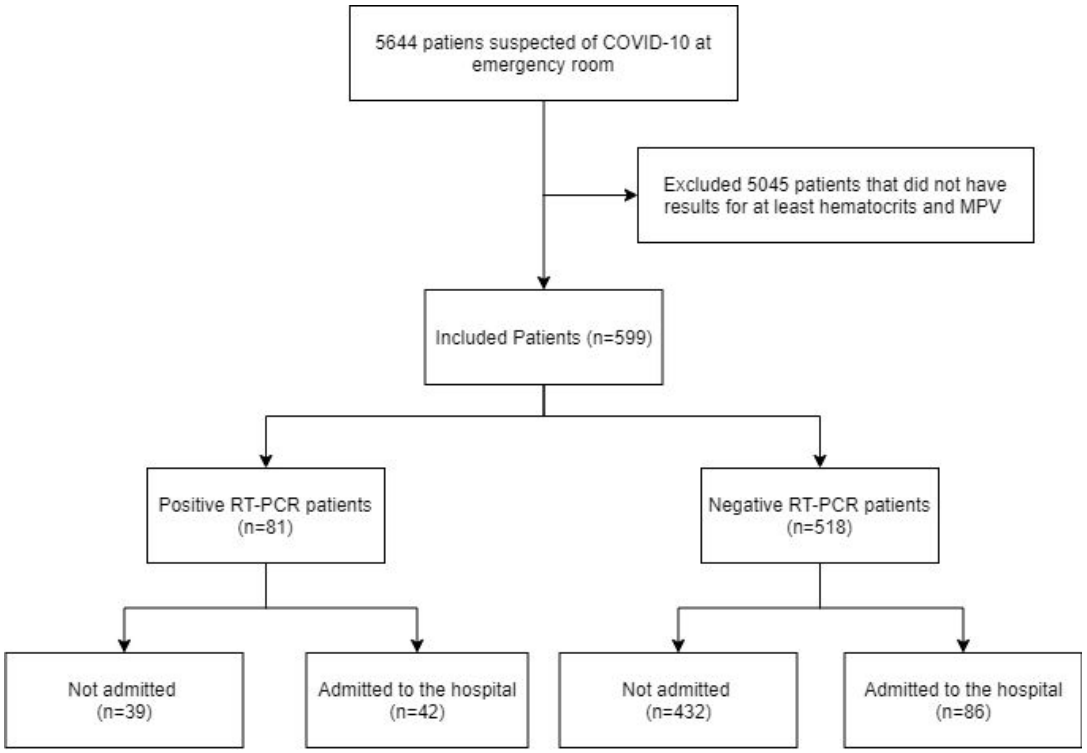

Figure 3 : Diagram representing the process to select the included patients in our study. Notice that from an initial number of 5,644 patients, we had adequate data for only 599.

Results for the artificial intelligence predictive model

We applied our proposed AI method to the already described dataset and repeated the training of the model and the evaluation of the metrics of interest in the test set 100 times, to guarantee that the results we report are not biased by a specific data partition. We also use the values of the replications to create the confidence intervals for each metric.

Reported Specificity was 92.16% [95% CI 91.73 - 92.59]; Sensitivity was 63.98% [95% CI 59.82 - 67.50], Negative Predictive Value (NPV) was 95.29% [95% CI 94.65% - 95.90%], whereas Positive Predictive Value (PPV) was 48.00% [95% CI 44.88% - 51.56%]. In Figures 4 to 7, we show the distribution of these metrics.

Table 1 presents the average contingency table taken as proportions over the 100 performed 100 replications of our test. One can notice an average of only 4% false negatives, which corroborates the findings of our study and the robustness of ER-CoV.

Table 1: Average contingency table shown as proportions for the 100 replications of ER-CoV

		Reference test (RT-PCR)		
		Positive	Negative	Total
Our test (ER-CoV)	Positive	6.50%	7.07%	13.57%
	Negative	4.06%	82.37%	86.43%
	Total	10.56%	89.444	100%

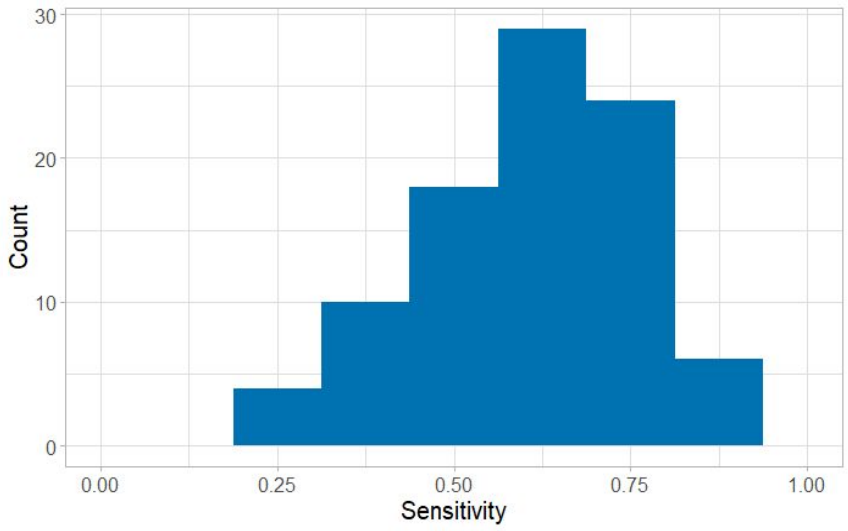

Figure 4: Histogram for the sensitivity of ER-CoV considering the 100 repetitions

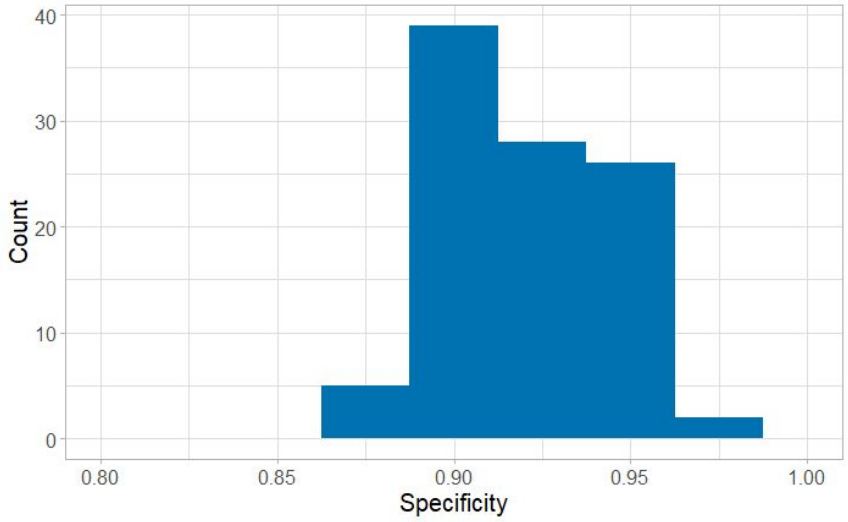

Figure 5: Histogram for the specificity of ER-CoV considering the 100 repetitions

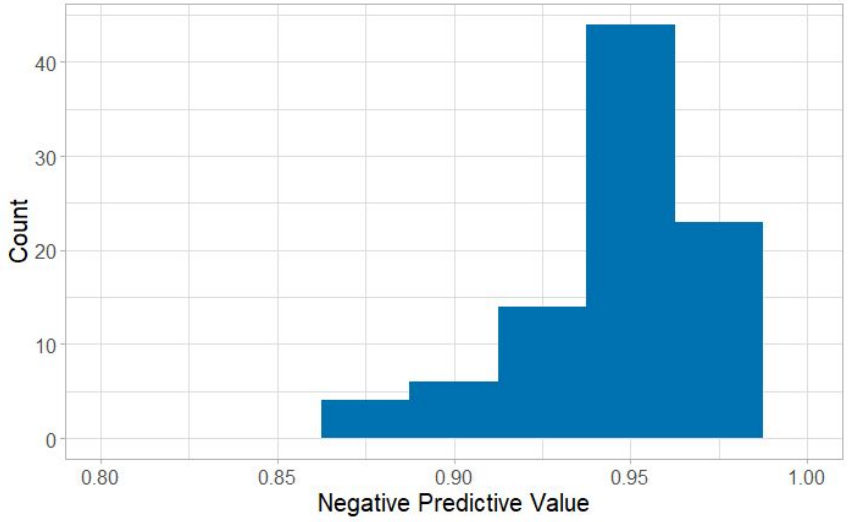

Figure 6: Histogram for the NPV of ER-CoV considering the 100 repetitions

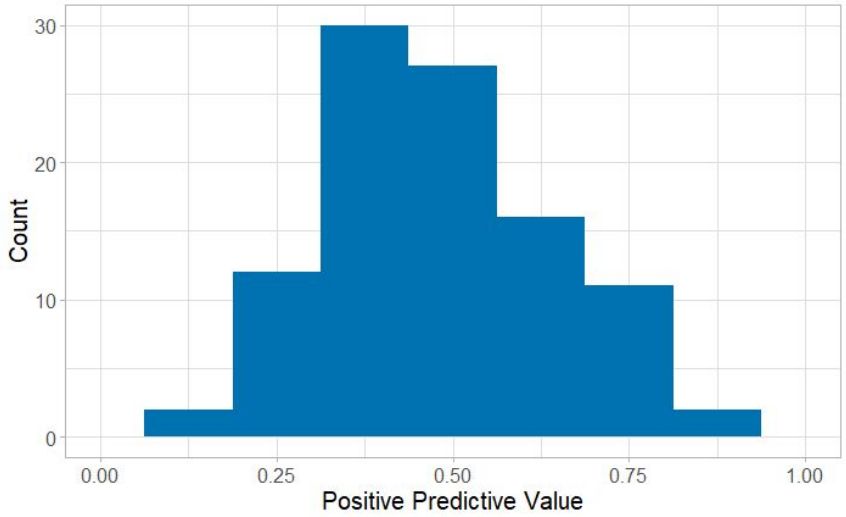

Figure 6: Histogram for the PPV of ER-CoV considering the 100 repetitions

We found that the proposed AI method was successful at discarding negative patients, while flagging potential positive patients for COVID-19. When performing error analysis of the 100 replications, we found that 99 out of 599 patients were misclassified at least once, as shown in Table 2.

Table 2: Summary of the number of unique patients that were misclassified during the repetitions of our experiments and their status regarding hospital admission

	RT-PCR Positive	RT-PCR Negative
Admitted	27	3
Non-Admitted	33	36
Total	60	39

45% of the patients misclassified as negative, that is false negatives, were admitted to hospital anyway, possibly due to the severity of their symptoms or potentially other factors such as comorbidities that would be available to the clinician justifying hospital admittance. For comparison, only 9.31% of the COVID-19 positive patients in the entire data set were hospitalised. Thus, even with an average sensitivity of 63.98% using only a limited number of blood exams (which already is an important finding), it is also relevant to take into account the severity of their symptoms and their clinical evaluation by a professional, with hospitalisation instead of going back to the community where they could transmit the infection.

Due to length limit, we added as supplementary material the results achieved using other setup scenarios, such as different training and test ratios and other classifiers. All results can be reproduced using the R Markdown file provided.

Discussion

By the simple introduction of the AI model, given its average specificity of 92.16%, the number of traditional tests already employed at the hospital's ER would be reduced by this factor. That is, only around 9% of suspected COVID-19 cases would needed to be sent for further testing. Due to its high Negative Predictive Value of 95.29%, less than 5% of patients would be misclassified as not being infected with COVID-19.

ER-CoV has the potential to improve COVID-19 screening at emergency rooms. The first, and possibly most important, benefit is the reduction of the number of tests required to be performed on patients that are negative for COVID-19.Another benefit is the potential to develop prioritisation queues for patients that our algorithm identifies as positive, thus

speeding up results for potentially infected patients. An additional positive impact is related to hospital and patient flow management. Hospitals would be able to provide better isolation of COVID-19 patients, given that blood samples could be drawn on a daily basis and ER-CoV would identify patients that should be moved to another ward, for instance a ward of suspected positive patients.

A limitation of our study is that we were not capable of identifying with high certainty which blood exams contribute the most to the classification due to the nature of our AI model framework. However, previous studies already identified that C reactive protein (Ling, 2020), leukocytes (Cheng et al., 2020; Qin et al., 2020), platelets (Cheng et al., 2020), and lymphocytes (Xu et al., 2020) are altered at different levels in COVID-19 patients. Thus, we envision for future research a detailed study of which blood exams are more informative for differential diagnosis, as well as for understanding how the SARS-CoV-2 virus alters blood components.

Another possible limitation is that data were collected only at the emergency room, with patients already displaying symptoms compatible with COVID-19. At this point, due to the lack of data from asymptomatic patients, we cannot generalize how our model would perform for a group of individuals that are not compatible with characteristic symptoms of COVID-19.

In this paper we report a novel method for the classification of COVID-19 patients at emergency rooms. ER-CoV is low-cost and relies only on simple blood exams that are fast and highly available, and resort to artificial intelligence methods to model such patients. We achieved extremely significant results and foresee many applications of this framework. Thus, we call for additional initiatives such as this one executed by the Albert Einstein Hospital, since data that can be easily anonymized could provide important insights in longitudinal studies of disease progression.

Acknowledgments

This work was supported by the Amazon AWS Cloud Credits for Research, the Google TensorFlow Research Cloud for the free credits for the usage of TPU. Their role was to provide computational resources for all tested models.

The University of Sheffield had the role of financially supporting the Ph.D scholarship of Felipe Soares.

Ethical Approval

We used only secondary anonymized data for our analyses. Self-declaration approval was given by the University of Sheffield Research Ethics Committee (Reference Number 034058).

The University Of Sheffield.

Downloaded: 08/04/2020
Approved: 08/04/2020

Aline Villavicencio
Computer Science

Dear Aline

PROJECT TITLE: A new high specificity COVID-19 screening method based on blood exams and artificial intelligence for initial triage in a constrained testing scenario
APPLICATION: Reference Number 034058

This letter confirms that you have signed a University Research Ethics Committee-approved self-declaration to confirm that your research will involve only existing research, clinical or other data that has been robustly anonymised. You have judged it to be unlikely that this project would cause offence to those who originally provided the data, should they become aware of it.

As such, on behalf of the University Research Ethics Committee, I can confirm that your project can go ahead on the basis of this self-declaration.

If during the course of the project you need to deviate significantly from the above-approved documentation please inform me since full ethical review may be required.

Yours sincerely

Com Ethics
Departmental Ethics Administrator

Declaration of Interest

All authors declare no conflict of interest

References

Attia, Z. I., Noseworthy, P. A., Lopez-Jimenez, F., Asirvatham, S. J., Deshmukh, A. J., Gersh, B. J., Carter, R. E., Yao, X., Rabinstein, A. A., Erickson, B. J., Kapa, S., & Friedman, P. A. (2019). An artificial intelligence-enabled ECG algorithm for the

identification of patients with atrial fibrillation during sinus rhythm: a retrospective analysis of outcome prediction. *The Lancet*, *394*(10201), 861–867.

Bahado-Singh, R. O., Yilmaz, A., Bisgin, H., Turkoglu, O., Kumar, P., Sherman, E., Mrazik, A., Odibo, A., & Graham, S. F. (2019). Artificial intelligence and the analysis of multi-platform metabolomics data for the detection of intrauterine growth restriction. *PloS One*, *14*(4), e0214121.

Boser, B. E., Guyon, I. M., & Vapnik, V. N. (1992). A training algorithm for optimal margin classifiers. *Proceedings of the Fifth Annual Workshop on Computational Learning Theory*, 144–152.

Chawla, N. V., Lazarevic, A., Hall, L. O., & Bowyer, K. W. (2003). SMOTEBoost: Improving Prediction of the Minority Class in Boosting. *Knowledge Discovery in Databases: PKDD 2003*, 107–119.

Cheng, Z., Lu, Y., Cao, Q., Qin, L., Pan, Z., Yan, F., & Yang, W. (2020). Clinical Features and Chest CT Manifestations of Coronavirus Disease 2019 (COVID-19) in a Single-Center Study in Shanghai, China. *AJR. American Journal of Roentgenology*, 1–6.

Corman, V. M., Landt, O., Kaiser, M., Molenkamp, R., Meijer, A., Chu, D. K. W., Bleicker, T., Brünink, S., Schneider, J., Schmidt, M. L., Mulders, D. G. J. C., Haagmans, B. L., van der Veer, B., van den Brink, S., Wijsman, L., Goderski, G., Romette, J.-L., Ellis, J., Zambon, M., … Drosten, C. (2020). Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. *Euro Surveillance: Bulletin Europeen Sur Les Maladies Transmissibles = European Communicable Disease Bulletin*, *25*(3). https://doi.org/10.2807/1560-7917.ES.2020.25.3.2000045

Efron, B. (1987). Better Bootstrap Confidence Intervals. *Journal of the American Statistical Association*, *82*(397), 171–185.

Kirkwood, B. R., & Sterne, J. A. C. (2010). *Essential Medical Statistics*. John Wiley & Sons.

Kumar, R., Sharma, A., Siddiqui, M. H., & Tiwari, R. K. (2018). Prediction of Drug-Plasma

Protein Binding Using Artificial Intelligence Based Algorithms. *Combinatorial Chemistry & High Throughput Screening*, *21*(1), 57–64.

Lake, M. A. (2020). What we know so far: COVID-19 current clinical knowledge and research. *Clinical Medicine*, *20*(2), 124–127.

Ling, W. (2020). C-reactive protein levels in the early stage of COVID-19. *Medecine et Maladies Infectieuses*. https://doi.org/10.1016/j.medmal.2020.03.007

Li, Z., Yi, Y., Luo, X., Xiong, N., Liu, Y., Li, S., Sun, R., Wang, Y., Hu, B., Chen, W., Zhang, Y., Wang, J., Huang, B., Lin, Y., Yang, J., Cai, W., Wang, X., Cheng, J., Chen, Z., … Ye, F. (2020). Development and Clinical Application of A Rapid IgM-IgG Combined Antibody Test for SARS-CoV-2 Infection Diagnosis. *Journal of Medical Virology*. https://doi.org/10.1002/jmv.25727

Phelan, A. L., Katz, R., & Gostin, L. O. (2020). The Novel Coronavirus Originating in Wuhan, China: Challenges for Global Health Governance. *JAMA: The Journal of the American Medical Association*. https://doi.org/10.1001/jama.2020.1097

Qin, C., Zhou, L., Hu, Z., Zhang, S., Yang, S., Tao, Y., Xie, C., Ma, K., Shang, K., Wang, W., & Tian, D.-S. (2020). Dysregulation of immune response in patients with COVID-19 in Wuhan, China. *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America*. https://doi.org/10.1093/cid/ciaa248

Rokach, L. (2010). Ensemble-based classifiers. *Artificial Intelligence Review*, *33*(1-2), 1–39.

Soares, F., Becker, K., & Anzanello, M. J. (2017). A hierarchical classifier based on human blood plasma fluorescence for non-invasive colorectal cancer screening. *Artificial Intelligence in Medicine*, *82*, 1–10.

Torgo, L. (2016). *Data Mining with R: Learning with Case Studies, Second Edition*. CRC Press.

Xu, Y.-H., Dong, J.-H., An, W.-M., Lv, X.-Y., Yin, X.-P., Zhang, J.-Z., Dong, L., Ma, X., Zhang, H.-J., & Gao, B.-L. (2020). Clinical and computed tomographic imaging features

of novel coronavirus pneumonia caused by SARS-CoV-2. *The Journal of Infection*, *80*(4), 394–400.