*J. Health Inform.* 2020 Número Especial SBIS - Dezembro: 358-65 **Artigo Original**

JHI JOURNAL OF HEALTH INFORMATICS

CBIS'20 7 a 11 de Dezembro Foz do Iguaçú - Paraná Evento Virtual

# A COVID-19 surveillance platform to monitor risk of infection based on a machine learning model

Plataforma de vigilância COVID-19 para monitorar o risco de infecção com base em um modelo de aprendizado de máquina

Plataforma de vigilancia COVID-19 para monitorear el riesgo de infección basada en un modelo de aprendizaje automático

**Daniel Mário de Lima[1], Ramon Alfredo Moreno[2], Marina de Sá Rebelo[3], José Eduardo Krieger[4], Marco Antonio Gutierrez[5]**

**Keywords:** Coronavirus Infections; Data Science; Machine Learning

## ABSTRACT

**Objective:** To develop a platform for daily survey of COVID-19 signs and symptoms in health employees to indicate the need of additional individual diagnostic procedures and to assist institutional planning to prevent the spread of the virus and sustain the hospital operations during the pandemic. **Methods:** We used information from a recent meta-analysis to simulate datasets of patients with different signs, symptoms and comorbidities to evaluate machine-learning algorithms for each dataset classification. The best performing model identifying COVID-19 from other similar conditions including H1N1 and seasonal influenza was selected as the base model for developing a platform for risk assessment. **Results and Conclusion:** The platform was deployed for surveillance of 4,200 collaborators from a tertiary hospital on a voluntary basis, but it can be readily adapted for other environments or populational surveillance to assist public authorities devising strategies to prevent the spread of the virus.

**Descritores:** Infecções por Coronavirus; Ciência de Dados; Aprendizado de Máquina

## RESUMO

**Objetivo:** Desenvolver uma plataforma para levantamento diário dos sinais e sintomas de COVID-19 em profissionais de saúde para indicar a necessidade de procedimentos diagnósticos individuais adicionais e auxiliar no planejamento institucional para prevenir a propagação do vírus e sustentar as operações do hospital durante a pandemia. **Métodos:** Usamos informações de uma meta-análise recente para simular conjuntos de dados de pacientes com diferentes sinais, sintomas e comorbidades para avaliar algoritmos de aprendizado de máquina para cada classificação de conjunto de dados. O modelo de melhor desempenho para identificar COVID-19 de outras condições semelhantes, incluindo H1N1 e influenza sazonal, foi selecionado no desenvolvimento de uma plataforma para avaliação de risco. **Resultados e conclusão:** A plataforma foi implantada para vigilância voluntária de 4.200 colaboradores de um hospital terciário, mas pode ser prontamente adaptada para outros ambientes ou vigilância populacional para auxiliar o poder público a traçar estratégias de prevenção à disseminação do vírus.

**Descriptores:** Infecciones por Coronavirus; Ciencia de los Datos; Aprendizaje Automático

## RESUMEN

**Objetivo:** Desarrollar una plataforma para la encuesta diaria de signos y síntomas de COVID-19 en el personal de salud para indicar la necesidad de procedimientos adicionales de diagnóstico individual y ayudar a la planificación institucional para prevenir la propagación del virus y mantener las operaciones del hospital durante la pandemia. **Métodos:** Usamos información de un metanálisis reciente para simular conjuntos de datos de pacientes con diferentes signos, síntomas y comorbilidades para evaluar algoritmos de aprendizaje automático para cada clasificación de conjuntos de datos. Se seleccionó el modelo de mejor desempeño para identificar COVID-19 a partir de otras afecciones similares, incluida la gripe H1N1 y estacional, como base para desarrollar una plataforma para la evaluación de riesgos. **Resultados y Conclusión:** La plataforma se implementó para la vigilancia de 4.200 colaboradores de un hospital terciario de forma voluntaria, pero puede adaptarse fácilmente a otros entornos o vigilancia poblacional para ayudar a las autoridades públicas a diseñar estrategias para prevenir la propagación del virus.

*[1] Mestre em Ciências da Computação, Laboratório de Informática Biomédica, InCor- Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo - HCFMUSP, São Paulo (SP), Brasil.*

*[2] Doutor em Engenharia Elétrica, Laboratório de Informática Biomédica, InCor-Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo - HCFMUSP, São Paulo (SP), Brasil.*

*[3] Doutor em Engenharia Biomédica, Laboratório de Informática Biomédica, InCor-Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo - HCFMUSP, São Paulo (SP), Brasil.*

*[4] Docente Titular de Genética e Medicina Molecular da Faculdade de Medicina-USP, São Paulo (SP), Brasil.*

*[5] Livre Docente da Faculdade de Medicina-USP, São Paulo (SP), Brasil.*

Autor Coorespondente: **Daniel Mário de Lima**
e-mail: **daniel.lima@hc.fm.usp.br**

## INTRODUCTION

In late 2019, a novel type of coronavirus, known as SARS-CoV-2, was discovered, causing several infections and pneumonia cases, first in Wuhan, China, with later spreading worldwide. The World Health Organization named the acute infectious disease caused by the SARS-CoV-2 as COVID-19 (Coronavirus Disease - 2019), a systemic infectious disease characterized by its high transmissibility.

A challenge in this type of emergency is the prompt detection of infected people to prioritize health care and prevent the spread of the virus, especially under low virus testing capacity, as is the case in many places including Brazil (depicted in Figure 1).

The lack of capacity for prompt diagnostics is particularly critical for healthcare workers (HCW) who have been overly affected and become an easy target for the infection during the pandemic[1]. According to worldwide publications, these workers represent 19% of all reported COVID-19 cases in Spain[2], 9% in Italy[3], 4.6% in Germany[4], and 3.8% in China[5]. These figures represent a higher incidence of the infection than observed in general population[6-7]. This outcome is a consequence of the HCW's exposure to infection from patients and daily interaction with fellow staff that may have been already infected and undiagnosed.

The high rate of infection of HCW poses additional risks, as stated by the Imperial College COVID-19 response team[1], "*Transmission to and potentially among this high-risk group compromises both their own health and may contribute to nosocomial spread within hospitals*".

Such a spread raises the issue of protecting vulnerable patients from a potentially infectious workforce, as social distancing is not possible while caring for patients. Moreover, HCW screening programs show positive influence on the team´s self-confidence, decreasing absenteeism, and reducing psychological problems[8]. Therefore, establishing strategies for protecting HCWs from the infection by the new coronavirus and the screening and surveillance of symptoms should be a priority to assist management of hospitals and healthcare facilities. Reliable data sources and analytic tools are essential for healthcare decision-makers to establish policies to mitigate the pandemic disruption. Often, surveys rely only on simple models based on the number of symptoms declared by the respondent[9-10] and have limited discrimination power to discriminate from common symptoms associated to other viral diseases, such as influenza and H1N1[11-12].

Screening of HCW for identifying the most likely employee at risk of infection by the SARS-CoV-2 was paramount considering the availability of other diagnostic tools. Thus, we proposed a platform to assess COVID-19 risk score based on a machine-learning model. COVIDuc (COVID under control) was then developed as an application based on a set of signs, symptoms and comorbidities that may affect the susceptibility or the intensity of the infection for daily surveys to assess the risk among HCW.

## METHODS

### Predictive model for COVID-10 risk from symptoms and conditions

We reviewed diagnostic studies of influenza viruses and SARS-CoV-2 for the affected populations' characteristics, the prevalence of symptoms, conditions, and laboratory results in severe and moderate risk groups[11-12]. Tang et al.[11] described the similar symptomatology of pandemic, seasonal, and no-influenza cases in a large prospective study with 2,683 patients. Ma et al.[12] performed a systematic review and meta-analysis of 20 studies, involving 19 Health Centers and 11 Multi-Centers, and with around 53,000 patients with COVID-19. The symptoms and conditions

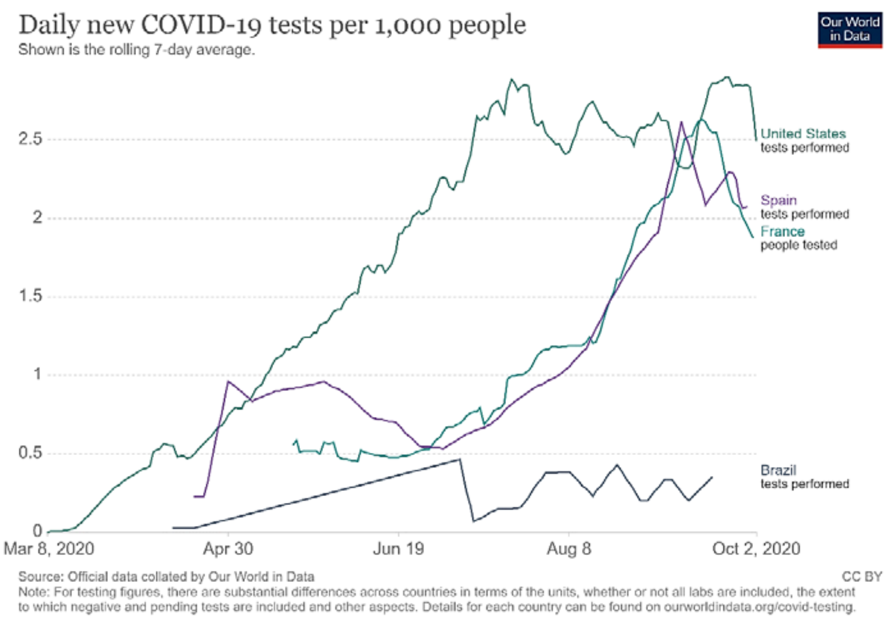



**Figure 1 –** The number of tests of COVID-19 by 1000 people according to "OurWorldInData.org" for Brazil, EUA, France, and Spain, from Mars to October 2020. (Source: https://ourworldindata.org/coronavirus#coronavirus-country-profiles)

found in these studies related to the disease were fever, cough, fatigue, expectoration, headache, diarrhea, myalgia, shortness of breath, sore throat/pharyngeal, nausea or vomiting, chill, nasal congestion/rhinorrhea, dyspnea, anorexia, dizziness, hypertension, diabetes, cardiovascular disease (CVD), cerebrovascular disease, chronic obstructive pulmonary disease (COPD).

Assuming symptoms and conditions as independent and normal variables, we simulated groups at different risk levels by random sampling of multivariate normal distributions, with similar prevalence of the reviewed clinical studies' symptoms and conditions. We prepared a data set comprising five groups with the following numbers of case studies: Group 1: 5000 simulating the "non-influenza"; Group 2: 5000 simulating the "seasonal influenza"; Group 3: 5000 simulating the "H1N1 (2009)"; Group 4: 7500 simulating the "COVID+ non-severe"; Group 5: 7500 simulating the "COVID+ severe" group. In this data set, 50% of the simulated case studies were positive COVID, and 50% were negative COVID. We then adjusted a linear support vector classifier (SVC)[13] to verify the "COVID+ severe" group's separability from the rest of the data. Following, we evaluated several machine learning (ML) algorithms for the classification of the groups. (1) BernoulliNB: a naïve Bayes model for multivariate Bernoulli distributions, that is, with binary-valued variables[14]. (2) DecisionTreeClassifier: if-then-else decision trees learned by splitting attributes' values into subsets of minimal (Gini index) classification error[15-16]. (3) LogisticRegression: the logit regression model for binary-valued output variables[17], adjusted by liblinear's dual coordinate-descent method[13]. (4) KNneighborsClassifier and RadiusNeighborsClassifier: indexed search of *k* neighbors in a kd-tree[18] or ball-tree[19], optionally limited to a radius *r* in the Minkowski metric space near the query point. (5) Random Tree Ensembles (RandomForestClassifier, ExtraTreesClassifier, AdaBoostClassifier): ensembles of randomized and bootstrapped decision trees[20-21], with random splitting thresholds[22], and gradient boosting[23].

We compared each classifier's predictive performance using ROC analysis at each new iteration of evaluation and adjustment of the symptoms and conditions list. Finally, we scaled the weights of the probability estimator of the logistic output to weights of a weighted average with an output between 0 and 10, which defines the range of the risk score for COVID-19. Figure 2 represents the principal steps involved in the modeling of the risk score for COVID-19.

The training was carried out with 80% of the data in 5-fold cross-validation, and the remaining 20% of the data were used to analyze the classifiers' operational characteristics of the reception curve (ROC curve). These analyses were executed in a Dell XPS8930 workstation with Intel i7-8700 CPU, 16GB RAM, 1TB HDD, and NVIDIA 1050Ti-4GB GPU; using Microsoft Excel, Python, Pandas, and Numpy for data preparation; Matplotlib for visualization; Scipy, StatsModels, and Patsy for statistical analyses; and Scikit-Learn for machine learning.

**Development and Deployment of the COVIDuc web application**

We have elaborated a questionnaire in collaboration with the hospital's clinical specialists to implement a pilot application for COVID-19 risk assessment of the hospital's employees. The survey comprised questions about a simplified subset of the symptoms and conditions used in the predictive model. We aggregated and filtered the new variables (the subset of symptoms and conditions) and trained the model using this simplified data set for the implementation of the online application prototype.

The probability of COVID-19 positive estimated by the model like logistic regression was adjusted as weights of a weighted average, with output between 0 and 10. After the revision performed by clinicians, and the re-training of the developed model with the reduced set of symptoms and conditions, we implemented a web application for risk assessment of the employees in a tertiary hospital, called COVID Under Control (COVIDuc). The implemented model takes into account the possible presence of the following pre-conditions and symptoms:

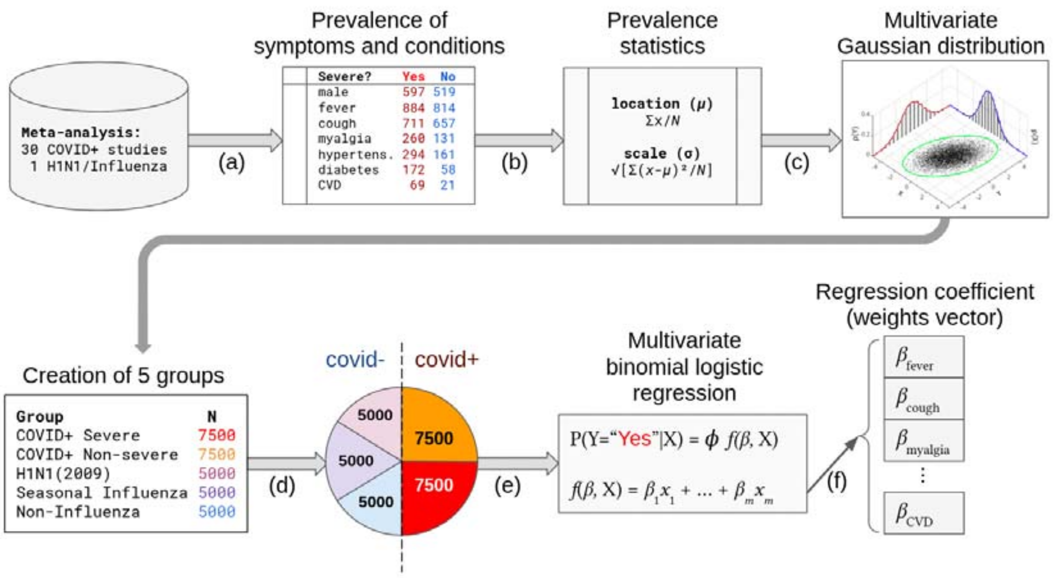



**Figure 2 –** Principal steps involved in the modeling of the risk score for COVID-19.

**1. Signals and Symptoms**: dry cough or phlegm; muscle aches; headache; sore throat; shortness of breath; stuffy or runny nose; diarrhea/nausea/vomit; chills or fever (> 38° Celsius); loss of taste or smell.

**2. Comorbidities:** hypertension; diabetes; heart disease; pulmonary disease; smoking.

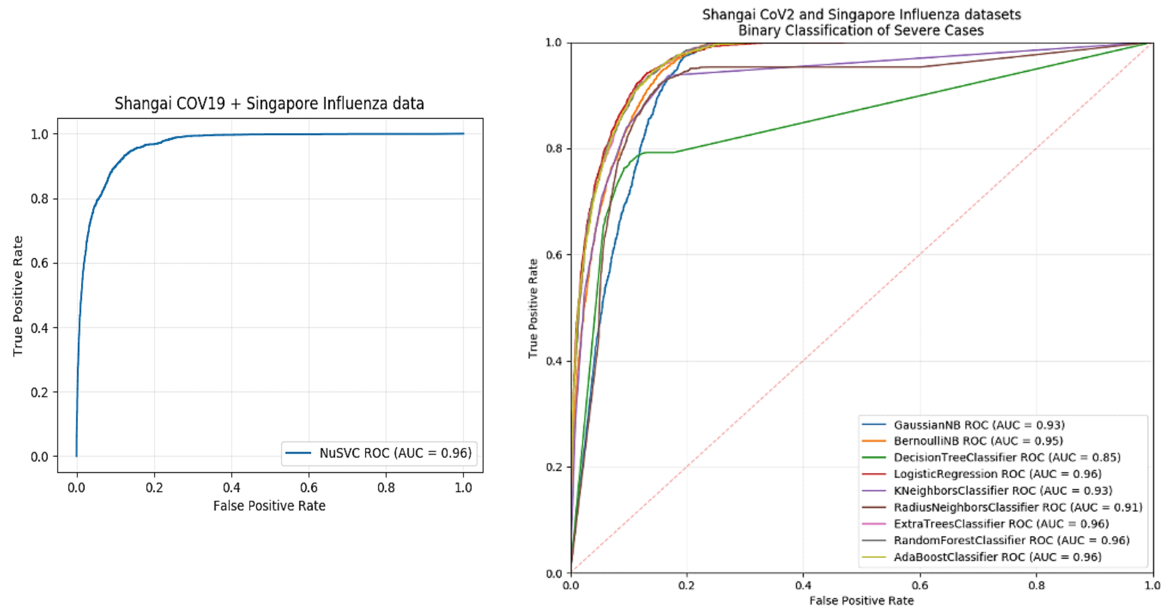
The frontend for the COVIDuc application was developed using Angular 10 and Enterprise Java Beans was used as backend. To monitor the use of the questionnaire by hospital´s employees and the resulting scores, we implemented two dashboards with metrics and statistics of accesses, with the application database in Oracle 12C.

## RESULTS

### Predictive model for COVID-10 risk from symptoms and conditions

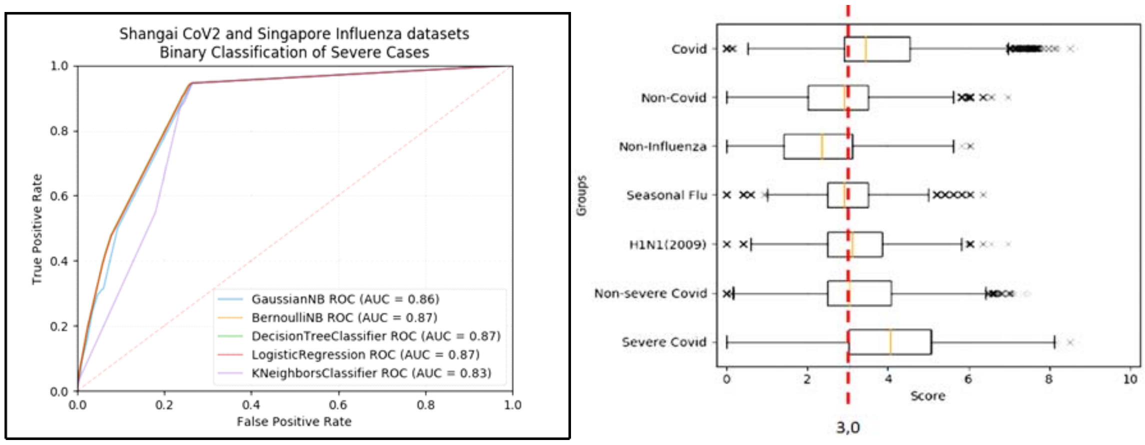
Figure 3 presents the ROC curves displaying the performance of the classifiers tested for the simulated groups with the complete set of symptoms from[11-12].

The ROC curves of the classification algorithms applied to the simulation groups (1 to 5) and the simplified data set are presented in Figure 4(a). After comparing the




(a) (b)

**Figure 3 –** (a) ROC analysis of the binary classification between “COVID+ Severe” and “other groups” using the “Nu Support Vector Classifier” algorithm, showing 96% of the area under the curve using data from Tang et al.[11] and Ma et al.[12]. (b) Comparative ROC analysis of binary classification algorithms between “COVID+ Severe” and “other groups”, using data from Tang et al.[11] and Ma et al.[12]. Logistic Regression, Random Forest, Extra Trees, and AdaBoost obtained the largest areas under the curve (96%).




(a) (b)

**Figure 4 –** (a) Comparative ROC analysis among binary classification algorithms between “COVID+ Severe” and “other groups” using the simplified set of symptoms and conditions, which are in the online questionnaire to be implemented. The largest area under the ROC curve was 87%. (b) Distribution of risk scores in each group (1 to 5). The boxes represent the median, interquartile range, 95% range, and extreme cases.

[I] https://qlik.com/products/qlik-sense performance of the classification algorithms, we decide to adopt the Logistic Regression whose coefficients define the weights for each symptom and condition.

Figure 4(b) presents the distribution of risk scores in the groups 1 to 5 and, after evaluation of this distribution, we set the risk score "3" as the threshold for risk of COVID-19 based on the set of symptoms and conditions from the chosen model's distribution of risk scores.

**Deployment of the COVIDuc web application**

To register in the COVIDuc, the user has to provide some personal information relevant to the application, such as age, weight, and height, and point out (if present) the symptoms and conditions. The COVIDuc application presents the resulting score, with the indication (or not) of the suspicion for COVID-19. After a few minutes, an SMS is sent to the user showing the risk score. Figure 5 shows the sequence of COVIDuc screens of the first access, from the initial page, with the consent form, to the result score and SMS.

Since the beginning of May, 4,200 collaborators from a tertiary hospital have been using the web application on a voluntary basis to assess their daily risk for COVID-19. The application has proven a valuable tool for HCW


**Figure 5** (In Portuguese) **–** The sequence of COVIDuc screens for the first access (in Portuguese). (1) From the QR code, the user can access the application. The QR code has been sent previously by email to all employees and is still available at multiple points at the hospital. (2) Disclaimer page. (3) Personal information page. (4) Screen with symptoms and pre-conditions. In the example screen, the person reported having diabetes as a pre-condition and the following symptoms: body pain, chills, or fever. (4) COVIDuc score page. For this example, there is a suspect of infection. Thus, the person is advised to go to the medical department for tests. (5) SMS is sent to the mobile phone number provided on the personal information page (6).

surveillance. The data collected is available for the managers at an Institutional dashboard. Figure 6 shows a screenshot of the dashboard, with daily access for all hospital areas and all risk ranges. The scores were available for the department of occupational medicine, which contacts the persons who have obtained high scores and have not shown up for health evaluation with a physician.

Figure 7 presents the number of people using the application, both daily and cumulative; they also present the number of persons that had any COVIDuc score greater or equal to 3.0, which was the adjusted threshold value of risk for COVID-19 by the developed model.

### Related Work

Early studies of COVID-19 data were focused on addressing the epidemiological aspects of the disease, the preparedness of hospitals for receiving patients, and the patient conditions in emergency care settings. The more general aspects of infectibility and disease spread are addressed by well-known epidemiological models, and are included in the WHO Global Influenza Surveillance and Response System (GISRS) protocols. Other studies analyze clinical parameters from readily available laboratory biomarkers, such as blood tests[24]. Another category of studies focused on active identification of suspects based on symptom assessment[25], which is alongside the COVIDuc objectives but for the general public instead. As all of the computational approaches require data for model development, many studies focus on preparing, integrating and publishing the clinical data for future retrospective studies[26].

## CONCLUSIONS AND FUTURE WORK

The model developed provides a method to assess the risk of Covid-19 based on the severity of signs and symptoms and conditions for COVID-19. We used data

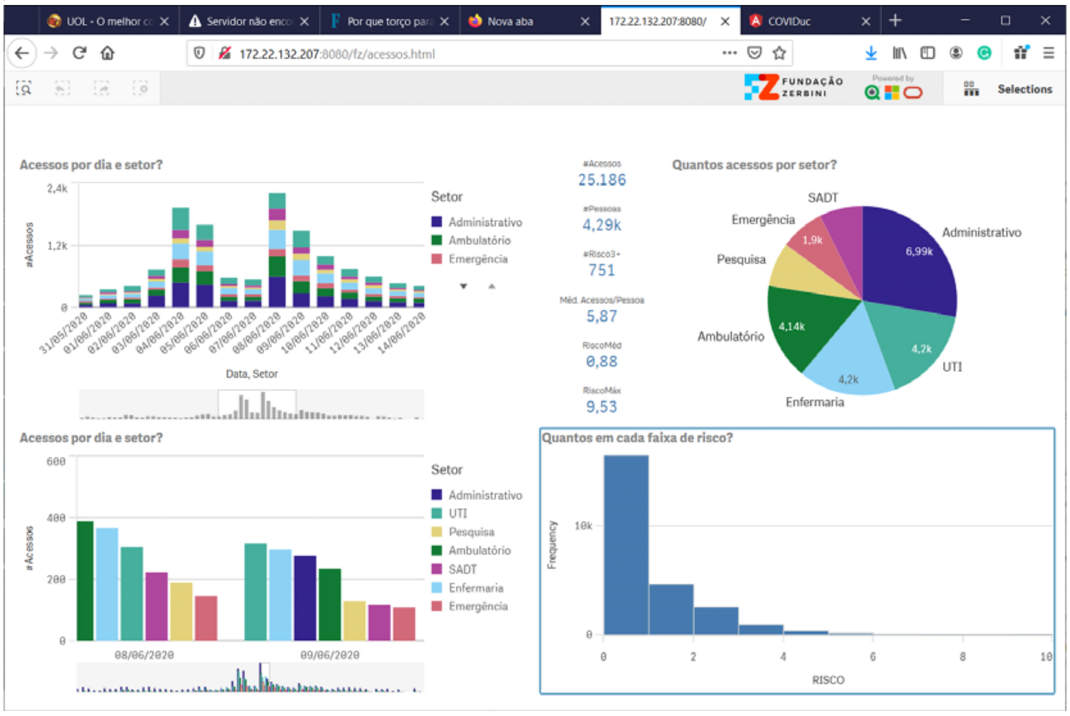
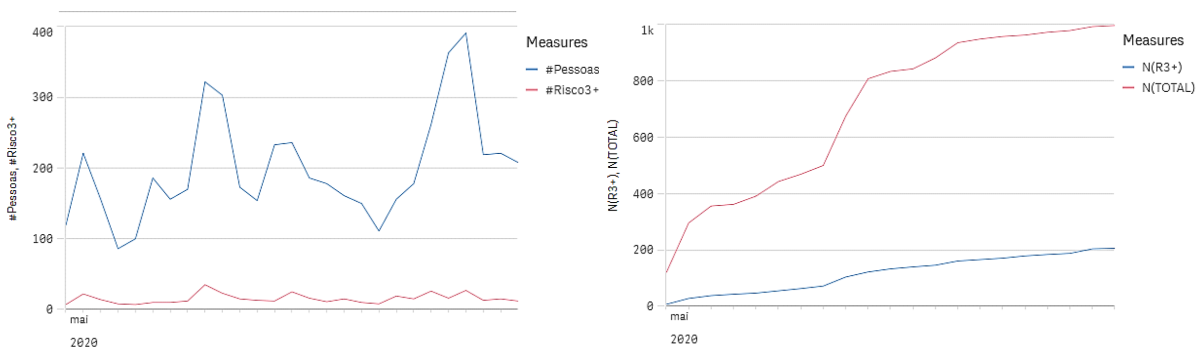



**Figure 6** (in Portuguese) **–** Dashboard presenting the data related to the respondents of the COVIDuc, daily access for area, how many persons in each risk range, among others.




(a) (b)

**Figure 7 –** (a) The daily number of people answering the questionnaire (blue line) and the number of persons with COVIDuc scores equal to or higher than 3.0 (red line), who were considered at risk for COVID-19 for May 2020. (b) The accumulated number of people answering the questionnaire (red line) and the number of persons with COVIDuc scores equal to or higher than 3.0 (blue line) since May 2020.

from a meta-analysis to define the weight of the factors contributing to influence the risk and this measure can be modified as definitive outcomes are verified by other specific diagnostics. The method does not use precise variables, such as laboratory and imaging exams and it is not meant for diagnostics, COVIDuc must be used as a tool for surveillance and screening only.

The collaborators from a tertiary hospital have used it as a surveillance tool since May 2020, on a voluntary basis, with the department of occupational medicine continually assessing the correspondence between the outcomes of the COVIDuc and the employee's medical files. COVIDuc was developed for daily surveillance of HCW, but it can be readily adapted for other environments or populational surveillance to assist public authorities devising strategies to prevent the spread of the virus. It was first used in a study carried out by the Solidary Research Network to evaluate public policies during the COVID-19 pandemic[13], and, since the middle of August, COVIDuc has been the screening tool at calling outpatients to resume the medical activities on the same tertiary hospital after the peak of the pandemic period in the city of Sao Paulo.

As future work, we will perform retrospective studies to test the models, as implemented at the COVIDuc application, using data obtained from the laboratory test for COVID-19 applied to the employees of the tertiary hospital. One of the possible studies is the comparison of the timeline of the COVIDuc scores with the laboratory results.

## ACKNOWLEDGMENTS

Our thanks go to Mr. João Henrique de Sá, who implemented a Web questionnaire suitable for gathering the data for assessing the COVID-19 risk.

Zerbini Foundation and Foxconn Brazil supported this work as part of the research project "Machine Learning in Cardiovascular Medicine". We also thank Dr. Diego Cardenas for his kind help in text revision.

## REFERÊNCIAS

1. Grassly N, Pons-Salort M, Parker E, White P, Ainslie K, Baguelin M, et al. Report 16: Role of testing in COVID-19 control. London, UK; 2020.
2. Ministerio de Sanidad España. Informe sobre la situación de COVID-19 en España no 22 [Internet]. 2020. Available from: https://www.isciii.es/QueHacemos/Servicios/VigilanciaSaludPublicaRENAVE/EnfermedadesTransmisibles/Documents/INFORMES/Informes COVID-19/Informe nº 22. Situación de COVID-19 en España a 13 de abril de 2020.pdf
3. International Council of Nurses. High proportion of healthcare workers with COVID-19 in Italy is a stark warning to the world: protecting nurses and their colleagues must be the number one priority [Internet]. 2020. Available from: https://www.icn.ch/sites/default/files/inline-files/PR_09_COVID-19 - Italy.pdf
4. Robert Koch Institute. Coronavirus Disease 2019 (COVID-19) Daily Situation Report of the Robert Koch Institute [Internet]. 2020. Available from: https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Situationsberichte/2020-04-13-en.pdf?__blob=publicationFile
5. Wu Z, McGoogan JM. Characteristics of and Important Lessons From the Coronavirus Disease 2019 (COVID-19) Outbreak in China: Summary of a Report of 72/ 314 Cases From the Chinese Center for Disease Control and Prevention. JAMA [Internet]. 2020 Apr 7;323(13):1239–42. Available from: https://doi.org/10.1001/jama.2020.2648
6. Pan A, Liu L, Wang C, Guo H, Hao X, Wang Q, et al. Association of Public Health Interventions With the Epidemiology of the COVID-19 Outbreak in Wuhan, China. JAMA [Internet]. 2020 May 19;323(19):1915–23. Available from: https://doi.org/10.1001/jama.2020.6130
7. Kluytmans-van den Bergh MFQ, Buiting AGM, Pas SD, Bentvelsen RG, van den Bijllaardt W, van Oudheusden AJG, et al. SARS-CoV-2 infection in 86 healthcare workers in two Dutch hospitals in March 2020. medRxiv [Internet]. 2020 Jan 1;2020.03.23.20041913. Available from: http://medrxiv.org/content/early/2020/03/ 31/2020.03.23.20041913.abstract
8. McAlonan GM, Lee AM, Cheung V, Cheung C, Tsang KWT, Sham PC, et al. Immediate and Sustained Psychological Impact of an Emerging Infectious Disease Outbreak on Health Care Workers. Can J Psychiatry [Internet]. 2007 Apr 1;52(4):241–7. Available from: https://doi.org/10.1177/070674370705200406
9. Rossman H, Keshet A, Shilo S, Gavrieli A, Bauman T, Cohen O, et al. A framework for identifying regional outbreak and spread of COVID-19 from one-minute population-wide surveys. Nat Med [Internet]. 2020;26(5):634–8. Available from: https://doi.org/10.1038/s41591-020-0857-9
10. Rivett L, Sridhar S, Sparkes D, Routledge M, Jones NK, Forrest S, et al. Screening of healthcare workers for SARS-CoV-2 highlights the role of asymptomatic carriage in COVID-19 transmission. Elife [Internet]. 2020;9(e58728). Available from: https://elifesciences.org/articles/58728
11. Tang JW-T, Tambyah PA, Lai FYL, Lee HK, Lee CK, Loh TP, et al. Differing Symptom Patterns in Early Pandemic vs Seasonal Influenza Infections. Arch Intern Med [Internet]. 2010 May 24;170(10):861–7. Available from: https://doi.org/ 10.1001/archinternmed.2010.108
12. Ma C, Gu J, Hou P, Zhang L, Bai Y, Guo Z, et al. Incidence, clinical characteristics and prognostic factor of patients with COVID-19: a systematic review and meta-analysis. medRxiv [Internet]. 2020 Jan 1;2020.03.17.20037572. Available from: http://medrxiv.org/content/early/2020/03/20/2020.03.17.20037572.abstract
13. Chou HY, Lin PY, Lin CJ. Dual coordinate-descent methods for linear one-class SVM and SVDD. In: Proceedings of the 2020 SIAM International Conference on Data Mining, SDM 2020. 2020. p. 181–9.
14. McCallum A, Nigam K. A comparison of event models for Naive Bayes text classification. In: ICML/AAAI-98 Workshop on Learning for Text Categorization [Internet]. 1998. p. 1–8. Available from: http://www.cs.cmu.edu/~knigam/papers/ multinomial-aaaiws98.pdf
15. Breiman L, Friedman JH, Olshen RA, Stone CJ. Classification and regression trees. Monterey, CA: Wadsworth & Brooks/Cole Advanced Books & Software; 1984. 358 p.
16. Quinlan JR. Induction of Decision Trees. Mach Learn. 1986;1(1):81–106.
17. Nelder JA, Wedderburn RWM. Generalized Linear Models. J R Stat Soc Ser A [Internet]. 1972 Jan 27;135(3):370–84. Available from: http://www.jstor.org/stable/ 2344614
18. Bentley JL. Multidimensional Binary Search Trees Used for Associative Searching. Commun ACM. 1975;18(9):509–17.
19. Omohundro SM. Five balltree construction algorithms. Vol. 51, Science. Berkeley, CA; 1989.
20. Breiman L. Bagging predictors. Mach Learn. 1996;24(2):123–40.
21. Breiman L. Random forests. Mach Learn. 2001;45(1):5–32.
22. Geurts P, Ernst D, Wehenkel L. Extremely randomized

trees. Mach Learn. 2006;63(1):3–42.
23. Freund Y, Schapire RE. A decision-theoretic generalization of on-line learning and an application to boosting. Lect Notes Comput Sci (including Subser Lect Notes Artif Intell Lect Notes Bioinformatics). 1995;904:23–37.
24. Batista AFM, Miraglia JL, Donato THR, Chiavegatto Filho ADP. COVID-19 diagnosis prediction in emergency care patients: a machine learning approach. medRxiv [Internet]. 2020 Jan 1;2020.04.04.20052092. Available from: http://medrxiv.org/content/early/2020/04/14/2020.04.04.20052092.abstract
25. Drew DA, Nguyen LH, Steves CJ, Menni C, Freydin M, Varsavsky T, et al. Rapid implementation of mobile technology for real-time epidemiology of COVID-19. Science [Internet]. 2020 Jun 19;368(6497):1362–1367. Available from: http://science.sciencemag.org/content/368/6497/1362.abstract
26. Mello LE, Suman A, Medeiros CB, Prado CA, Rizzatti EG, Nunes FLS, et al. Opening Brazilian COVID-19 patient data to support world research on pandemics [Internet]. Zenodo; 2020. Available from: https://doi.org/10.5281/zenodo.3966427