BMJ Open

Chronic lung lesions in COVID-19 survivors: predictive clinical model

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2021-059110
Article Type:	Original research
Date Submitted by the Author:	11-Nov-2021
Complete List of Authors:	Carvalho, Carlos; Universidade de São Paulo, Instituto do Coração - Divisão de Pneumologia Chate, Rodrigo; Universidade de São Paulo Hospital das Clínicas, Instituto de Radiologia Sawamura, Marcio; Universidade de Sao Paulo Hospital das Clinicas, Instituto de Radiologia Garcia, Michelle; Universidade de São Paulo Hospital das Clínicas, Instituto do Coração - Divisão de Pneumologia Lamas, Celina ; Universidade de Sao Paulo Hospital das Clinicas, Instituto do Coração - Divisão de Pneumologia Cardenas, Diego; Universidade de São Paulo Hospital das Clínicas, Instituto do Coração - Divisão de Informática Lima, Daniel; Universidade de São Paulo Hospital das Clínicas, Instituto do Coração - Divisão de Informática Scudeller, Paula; Universidade de São Paulo Hospital das Clínicas, Instituto do Coração - Divisão de Pneumologia Salge, João; Universidade de São Paulo Hospital das Clínicas, Instituto do Coração - Divisão de Pneumologia Nomura, Cesar; Universidade de São Paulo Hospital das Clínicas, Instituto de Radiologia Gutierrez, Marco; Universidade de São Paulo Hospital das Clínicas, Instituto do Coração - Divisão de Pneumologia
Keywords:	COVID-19, Chest imaging < RADIOLOGY & IMAGING, RESPIRATORY MEDICINE (see Thoracic Medicine)

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Chronic lung lesions in COVID-19 survivors: predictive clinical model

Carlos R R Carvalho (0000-0002-1618-8509)[1], Rodrigo C Chate[2], Marcio VY Sawamura[2], Michelle L Garcia[1], Celina A Lamas[1], Diego AC Cardenas[3], Daniel M Lima[3], Paula G Scudeller[1], João M Salge[1], Cesar H Nomura[2], Marco A Gutierrez[3], HCFMUSP Covid-19 Study Group*

1 Pulmonary Division, Heart Institute (InCor), Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo (HCFMUSP), Sao Paulo, SP, Brazil.

2 Radiology Institute (InRad), Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo (HCFMUSP), Sao Paulo, SP, Brazil.

3 Informatics Division, Heart Institute (InCor), Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo (HCFMUSP), Sao Paulo, SP, Brazil.

Correspondence to: Dr Carlos RR Carvalho (ORCI ID: 0000-0002-1618-8509), Pulmonary Division, Heart Institute (InCor), Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo (HCFMUSP), Av. Dr Eneas Carvalho de Aguiar, 44, Cerqueira Cesar, São Paulo, SP, 05403-900. Email: carlos.carvalho@hc.fm.usp.br. Fone: +55 11 26614505.

Word Count: 2872

Abstract

Objectives SARS-CoV-2 infection has many sequelae, including "fibrotic-like" lung lesions and lung function alterations, that will represent one of the major public health issues worldwide. Our objective was to propose a simple, accessible and low-cost predictive clinical model to detect lung lesions due to COVID-19 infection.

Design, settings and participants: This prospective cohort study included COVID-19 survivors hospitalized between March 30 and August 31, 2020, and re-examined after 6 months of hospital discharge from the ward or intensive care unit of a tertiary hospital (Hospital das Clínicas, Universidade de São Paulo), in Brazil. 749 patients (median [IQR] age, 56 [44.4-65.1] years; 53% male) followed the inclusion criteria (≥ 18 years patients with RT-PCR-confirmed SARS-CoV-2 infection) and were eligible for this study.

Outcome Measures: Demographic and anthropometric data were collected during interviews, and pulmonary function was assessed using the modified Medical Research Council(mMRC) dyspnoea scale, oximetry(SpO_2), spirometry(forced vital capacity[FVC]), and chest X-ray(CXR). Patients with changes in at least one of these examinations were invited to undergo chest computed tomography(CT). The results of mMRC scale, SpO_2, FVC, and CXR were used to train a machine learning model to detect lung lesions on CT.

Results After a general assessment, 470 patients (63%) presented at least one sign of pulmonary involvement and underwent CT. Among these, 48% had significant pulmonary changes, including ground-glass opacities, parenchymal bands, reticulation, traction bronchiectasis, and architectural distortion. The machine learning model accurately detected pulmonary lesions by the joint analysis of CXR, mMRC scale, SpO_2, and FVC data (Sensitivity [0.85±0.08], Specificity [0.70±0.06], F1-score [0.79±0.06] and AUC [0.80±0.07]).

Conclusion A predictive clinical model using CXR, mMRC, oximetry, and spirometry data can accurately screen patients with chronic lung lesions after SARS-CoV-2 infection. Given that these examinations are highly accessible and low cost, this protocol can be automated and implemented in different countries.

Strengths and limitations of this study

- Our study proposes a strategic tool for the identification of post-COVID patients with chronic lung lesion, which will represent one of the major public health issues worldwide.

- This study assessed in person the respiratory function in a large cohort of 749 critically or moderately ill COVID-19 patients that survived, covering a clinical, functional, and radiological aspects, while in previous studies most of the information was collected remotely, and pulmonary function was assessed in person in a few cases.

- 59% of patients from our cohort (N=445) was critically ill patients from ICU while few data on the pulmonary function of critically ill patients are available.

- The cohort population was heterogeneous and came from all districts of the metropolitan region of Sao Paulo, besides the single-centre nature of the study.

- The predictive clinical model proposed herein could guide countries at different levels of development to determine the treatment course in an early, fast and effective way, using accessible and low-cost examinations, besides reducing the radiation exposure.

INTRODUCTION

Coronavirus disease 2019 (COVID-19), caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), emerged in December 2019 and spread globally[1]. This multisystemic viral disease promotes endothelial and microvascular damage and immune system dysregulation, leading to hyperinflammatory and hypercoagulable states [2 3]. Several organs can be affected during the acute phase of COVID-19. The pulmonary complications are considered life-threatening because of the possibility of progressing to respiratory failure [4 5]. The scientific literature have already described that COVID-19 symptoms can persist for more than 12 weeks after acute infection, characterizing long COVID [1]. The clinical complains of dyspnoea, fatigue, cough, chest pain, depression, cognitive disorders, headache, palpitations, myalgia, and arthralgia are the most reported in long COVID[6-9]. In addition to symptoms, some studies have shown that radiological abnormalities are also frequent in the follow-up of patients after the acute phase. In one of them, chest computed tomography (CT) was performed in 171 patients 4 months after hospital discharge and showed abnormalities in 75.5% of the patients who required invasive mechanical ventilation (IMV)[10]. “Fibrotic-like changes” were observed in 19.3% of the total cohort and in 38.8% of patients with acute respiratory distress syndrome [9]. IMV can predict pulmonary sequelae, which reduce functional capacity and the health-related quality of life [6 11 12]. National Institute for Health and Care Excellence (NICE), have reported that some examinations can guide the diagnosis and management of post-COVID-19 syndrome [1], including oximetry, spirometry, chest X-ray (CXR), ultrasonography, modified Medical Research Council (mMRC) dyspnoea scale, and chest CT. The latter examination is the gold standard for the diagnosis of chronic lung lesions due to COVID-19 and characterization of “fibrotic-like” lung lesions [1 10].

The World Health Organization reported that more than 221 million COVID-19 cases were confirmed worldwide, with more than 4 million deaths, and more than 233 million patients recovered by September 2021 [13]. The large number of individuals with long-term symptoms has drawn the attention of several

countries[14 15]. For instance, in early 2021, the United Kingdom National Institute for Health Research invested £18.5 million to found Long COVID studies [16]. The lack of knowledge and medical training for treating post-COVID symptoms represents a significant public health challenge worldwide [14]. A reorganization of health systems will be necessary to address this issue, requiring the reallocation of resources and training of multidisciplinary teams and the development of comprehensive strategies and new approaches [14]. In this context, the wide availability of CRX and CT scanners has enabled the development of deep learning (DL) artificial intelligence-based algorithms for the automated diagnosis and prognosis of COVID-19 [17-19]. In an initiative, Castiglioni et al. [17] proposed a DL model for diagnosing COVID-19 with high sensitivity and specificity using radiography, while Wang et al. [18] developed a DL model (DenseNet) to classify CT images as positive or negative for COVID-19.

However, a more comprehensive protocol for screening COVID-19 patients and assessing the risk of chronic pulmonary changes in recovered patients has not been validated to date. Thus, the aim of this study was to develop a simple and accessible machine learning (ML)-based diagnostic protocol using the mMRC dyspnoea scale, oximetry, spirometry (forced vital capacity [FVC]), and CXR to detect the presence of radiologic chronic lung lesions due to SARS-CoV-2 infections.

METHODS

Study design and eligibility

This prospective cohort study detected chronic lung lesions in adult patients (≥ 18 years) with RT-PCR-confirmed SARS-CoV-2 infection admitted to the ward or intensive care unit (ICU) of the Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo (HCFMUSP), Sao Paulo, Brazil, from March 30 to August 31st, 2020. The protocols used in this study were described previously [20]. All research procedures were approved by the Research Ethics Committee of our institution (Process No. 31942020.0.000.0068).

The patients were invited to participate in the study six months after admission, and a face-to-face consultation was scheduled. Clinical, radiological, and laboratory evaluations were performed after the patients gave written

informed consent. Clinical data were stored in a structured form developed using REDCap software (https://www.redcapbrasil.com.br/).

General evaluation

Clinical data (comorbidities, cardiorespiratory symptoms, and smoking history), including the length of ICU stay and the need for IMV, were collected during semi-structured interviews. Anthropometric data and vital signs were also collected.

Pulmonary assessment was performed with an emphasis on respiratory symptoms. Dyspnoea was assessed using the mMRC scale [20]. Oxygen saturation (SpO_2) at rest and after physical exertion (1-min sit and stand test) was measured by pulse oximetry [20 21]. Spirometry was performed according to criteria established by ATS/ERS Task Force [22]. Actual spirometry results were compared with predicted values, according to Pereira et al. [23].

CXR was performed in posteroanterior and lateral views, according to standard guidelines. The results of the examinations were evaluated independently by two chest radiologists (MVYS and RCC, with 7 and 16 years of experience in thoracic radiology, respectively) working on dedicated workstations. The radiographs were scored as 0 (results were normal or not related to COVID-19 [including cardiomegaly and pulmonary nodules, for instance]) or 1 (findings which could be related to COVID-19 [including bilateral linear and/or reticular opacities, especially peripheral opacities]). Disagreements were resolved by consensus.

Previous classifications of radiographs were used to train and validate a DL algorithm with an EfficientNetB7 architecture [19]. A 5-fold cross-validation strategy was adopted for model training and validation, leading to an average area under the curve (AUC) of 0.89 (Supplementary Methods).

Chest CT

Patients who showed abnormalities during the initial assessment were enrolled to perform CT. The following criteria were used: (a) mMRC≥2; (b) resting $SpO_2 \leq 90\%$ and/or a decrease in SpO_2 of ≥4% during the 1-min sit and stand

test; (c) opacities likely related to COVID-19 on CXR; (d) FVC < lower limit of normal (LLN). The mean interval between CXR and chest CT was 45 ± 33 days.

The CT protocol used in this study was described previously [20]. CT findings consistent with COVID-19 were categorized according to the criteria of the Fleischner Society [24], including ground-glass and peripheral opacities, consolidations, parenchymal bands, reticulations, traction bronchiectasis, architectural distortions, honeycombing, bronchial wall thickening, mosaic attenuation, and pleural effusion.

The extent of lung involvement was quantified according to Francone et al. [25] by assigning the following scores to each pulmonary lobe: 0, none; 1, <5%; 2, 5-25%; 3, 26-50%; 4, 51-75%; 5, >75%. The total score varied from 0 to 25 and was calculated by summing the scores of the five lobes.

A score ≥7 was used as the cut off value for significant CT changes after model calibration. The equations used to determine these scores are described in the Supplementary Methods.

Machine learning (ML) model

A logistic regression-based ML model was used to detect the presence of COVID-19-related chronic lung lesions. In this model, the results of the mMRC scale, oximetry, and spirometry, and DL-based classification of CXR images were used as input data, and the presence of pulmonary lesions was used as output data (Figure 1).

Statistical analysis

Normally-distributed continuous variables were expressed as means and standard deviations, or medians and interquartile ranges. Categorical variables were compared using the chi-square test. Normally and non-normally distributed continuous variables were compared using Student's t-test and non-parametric tests, respectively (Excel 2016; Python 3.8.11; extension packages: Pandas 1.0.1; Numpy 1.19.5; Scipy 1.5.4; Scikit-Learn 0.24.0).

The performance of the DL model was assessed by the area under the receiver operating characteristic (AUC) curve, and the performance of the ML model was determined by the metrics Sensitivity, Specificity, F1-score and AUC (Supplementary Methods).

Patient and public involvement

Patients or the public were not involved in the design, conduct, reporting or dissemination plans of this research.

RESULTS

Of 3,753 enrolled patients, 1,957 were eligible for the study. Of these, 749 were included in the final analysis (445 [59%] and 304 [415] were admitted to the ICU and ward, respectively). Additional information on inclusion and exclusion criteria is shown in Figure 2.

Demographic data are shown in Supplementary Table S1. The median age of the cohort was 56 years, with a predominance of overweight individuals, and 53% were male. In our cohort, 59.4% of patients were admitted to the ICU; of these, 68.5% were on IMV during the study period. The vital signs of most patients were within normal limits (Supplementary Table S1).

The median interval between hospital admission and consultation was 7.1 (6.7–8.5) months, and the lower and upper limits were 5.4 and 12.9 months, respectively. Of a total of 749 patients, 470 (63%) had at least one sign of pulmonary involvement (Table 1). The Supplementary Figure S1, illustrates the simultaneous presence of two or more criteria for pulmonary involvement.

The demographic and clinical data regarding patients with or without a pulmonary involvement are described in Supplementary Table S2. Patients with pulmonary involvement were older and predominantly female. In addition, the number of comorbidities and the rate of ICU admission were higher in this population (Supplementary Table S2). In this group, 348 underwent CT (68%) (Figure 2). Demographic and clinical characteristics were similar between patients that underwent or did not undergo the CT (Supplementary Table S3).

CT scores were obtained from 328 (94%) patients. Scores were not determined in 20 patients because low image quality did not allow accurately assessing pulmonary changes. Chest CT analysis showed that 47.6% of the patients had a score ≥7, and the most common features were ground-glass opacities, parenchymal bands, reticulation, traction bronchiectasis, and architectural distortions (Supplementary Table S4). In this group, 86.5% and 13.5% were admitted to the ICU and ward, respectively. Among the patients with normal CT (score = 0), 36.4% and 63.6% were admitted to the ICU and ward, respectively. The frequency of CT changes is shown in Supplementary Table S5. The frequency of "fibrotic-like" lesions, including traction bronchiectasis and architectural distortion, was significantly higher in the group admitted to the ICU in the acute phase of the disease. Long-term CT features in patients with moderate and critical COVID-19 are shown in Figure 3 and Supplementary Figure S2, respectively.

Of 348 enrolees with CT data, 257 patients with results for mMRC, oximetry, spirometry, X-ray, and chest CT were selected for the prediction of pulmonary changes analysis. These changes were not assessed in 91 patients, since 61 cases did not present the results of all four tests (mMRC, oximetry, spirometry, CRX and CT) and 30 cases showed radiographic signs not related to COVID-19 (30 cases) (Supplementary Table S6).

The predictive performance of the ML model was evaluated by the metrics Sensitivity, Specificity, F1-score and AUC. A 5-fold cross-validation strategy was adopted for model training and validation. Three data groups were considered: (1) clinical data (oximetry [SpO_2], mMRC dyspnoea scores, and spirometry [FVC]), (2) CXR, and (3) all results (oximetry [SpO_2], mMRC dyspnoea scores, spirometry [FVC], and CXR). The performance of the predictive model was higher using the combination of all variables (clinical variables and CXR) considering the metrics Sensitivity of 0.85±0.08 (95% CI [0.77,0.94]), Specificity of 0.70±0.14 (95% CI [0.55, 0.85]), F1-score of 0.79±0.06 (95% CI [0.73, 0.85]), and AUC of 0.80±0.07(95% CI [0.72, 0.87]), expressed in terms of mean and standard deviation (Table 2).

Using the LR model, the predictive model is represented by the following function:

$$p_{CT} = 0.59 \times \left(\frac{FVC_{Resting}}{2FVC_{lln}}\right) - 2.16 \times \left(\frac{mMRC}{4}\right) + 0{,}679(SpO_2) + 1.15 \times p_{RX0} + 1.41 \times p_{RX1} + 1.04 \times p_{RX2} + 0.69 \times p_{RX3} + 0.60 \times p_{RX4}$$

where p_{CT} is the presence of abnormalities on CT images.

DISCUSSION

Few studies have assessed pulmonary changes in COVID-19 survivors after 6 months of hospital discharge. However, some of these patients developed long-term pulmonary complications after discharge [6 26-30]. The present study evaluated 749 patients who received supplemental oxygen or ventilatory support in the ward or ICU and survived. The cohort studied herein underwent an in person comprehensive clinical, functional, and radiological assessment, being extensive compared to previous studies in the literature[6 27 28 30-32], which confers reliability to this research.

In the first months after recovery, the most common CT changes previously described were ground-glass opacities, parenchymal bands, reticulation, mosaic attenuation pattern, and "fibrotic-like" features, including traction bronchiectasis and architectural distortions [33 34]. These alterations were present in 76.5% of our cohort, and severe and extensive changes were found in approximately 50% of the cases. The number of CT changes was higher in older critical patients, and individuals with more comorbidities, as well as was verified in other studies [29 35]. This result indicates that the prevalence of chronic lung lesions and sequelae in the post-COVID may be high worldwide. In this context, the development of strategies to deal with this issue will be necessary since the increased frequency of symptoms such as fatigue, weakness, and dyspnoea, and the presence of long-term complications impose a significant health and economic burden [14].

Therefore, the identification of severe pulmonary complications due to COVID-19, including fibrosis[1], and the large number of COVID-19 survivors, prompted us to developed a predictive clinical model to screen patients admitted

to a tertiary hospital to reduce costs and radiation exposure. During the first 6 months of the pandemic in Sao Paulo, Brazil, all hospital beds at HCFMUSP (300 in the ICU and 400 in the ward) were made available to COVID-19 patients [12]. Patients are treated free of charge in our hospital in a universal health system, and there is a constant search for better and cost-effective protocols to improve workflow [12]. Thus, we demonstrated herein the possibility of using a protocol involving simple and accessible examinations, such as the mMRC dyspnoea scale, oximetry, spirometry, and CXR.

Dyspnoea scales, CXR, oximetry, and spirometry are commonly used to evaluate COVID-19 symptoms [2]. A Norwegian study evaluated a cohort of 100 patients 3 months after hospital admission and showed that 19% had dyspnoea (mMRC scores>1), and 10% presented altered FVC and normal oxygen saturation, suggesting that the sensitivity of pulse oximetry is lower [36]. In 113 patients evaluated 4 months after COVID-19 diagnosis, FVC and oxygen saturation were lower in severe cases than in moderate cases, although the mean values remained within the limits of normality [32]. In addition, a previous study has pointed that cough, lymphocytosis and the lung volume could indicate lung lesions in COVID-19 recovered patients [31].

Ground-glass and reticular opacities can be detected by CXR, despite being less sensitive than CT (25). In addition, CXR is readily available in the primary care setting and has a lower cost and level of radiation than CT (25). Radiographs were scored by an automated DL-based image analysis tool and by chest specialists, and there was a high level of consensus between these strategies (AUC of 0.89). In the Brazilian public health system, the cost of a CT scan is approximately 15 times higher than that of a CRX [37]. The American College of Radiology and the Radiological Society of North American show that radiation doses of a standard chest CT and CXR are 6.1 mSv and 0.1 mSv, respectively, underscoring the possibility of reducing exposure to ionizing radiation, especially in a population serially exposed to imaging procedures in the acute phase of COVID-19 [38].

Nevertheless, none of these examinations by themselves accurately predicted pulmonary complications. The performance of our model corroborates

this finding since the information provided by each clinical examination alone did not accurately diagnose the pulmonary changes detected on CT. In contrast, clinical and radiographic data were complementary and increased the performance of the ML model. Furthermore, cross-validation increased the robustness of the results. These results indicate that four examinations (oximetry, mMRC dyspnoea scale, spirometry, and CXR) should be jointly executed to screen patients at risk of developing chronic lung lesions due to COVID-19 and achieve a diagnostic performance similar to that of CT (Sensitivity of 0.85±0.08, Specificity of 0.70±0.14, F1-score of 0.79±0.06 and AUC of 0.80±0.07). The analysis of these metrics indicates that the method can better identify the true positives when compared to the ability to identify the true negatives. In addition, the F1-score takes into account false positives and false negatives and measures the accuracy of the method in the dataset.

Our study has limitations. First, there was variability in the interval between the execution of CXR and CT. Notwithstanding this variation, which might contribute to lung recovery, our protocol screened a large number of patients with pulmonary lesions, demonstrating the persistence of these manifestations secondary to COVID-19 and reducing sampling bias. Second, the single-centre nature of the study limits the generalizability of the results. However, a previous study showed that the population of patients admitted to HCFMUSP—a tertiary reference hospital for the treatment of COVID-19 in Brazil—was heterogeneous and came from all districts of the metropolitan region of Sao Paulo [12]. Third, we were unable to contact some patients because of inconsistencies in telephone numbers and addresses. Thus, these subjects were not included in the protocol, although public death registry data showed that they were alive. Fourth, this screening protocol was developed based on respiratory complaints, which are considered risk factors for developing chronic lung complications. However, other COVID-19 symptoms were not analysed in this study.

The breadth of our results allowed us to propose a simple, accessible, and low-cost clinical predictive model to screen patients at risk of developing chronic lung lesions due to COVID-19. The low cost and easy access to these examinations allow implementing this protocol in developing countries. Also, it enables to determine the treatment course in an early, fast and effective way,

reducing the radiation exposure as well as the execution time and cost of imaging examinations. The use of artificial intelligence allowed the large-scale assessment of radiographs and their association with clinical, demonstrating that artificial intelligence models can be used to automate diagnosis, especially in severe patients.

Colaborators: *Members of the HCFMUSP Covid-19 Study Group: Adriana L Araújo, Aluisio C Segurado, Amanda C Montal, Anna Miethke-Morais, Anna S Levin, Beatriz Perondi, Bruno F Guedes, Carolina Carmo, Carolina S Lázari, Cassiano C Antonio, Clarice Tanaka, Claudia C Leite, Cristiano Gomes, Edivaldo M Utiyama, Emmanuel A Burdmann, Eloisa Bonfá, Esper G Kallas, Ester Sabino, Euripedes C Miguel, Fabio R Pinna, Fabiane Y O Kawano, Geraldo F Busatto, Giovanni G Cerri, Guilherme Fonseca, Heraldo P Souza, Izabel Marcilio, Izabel C Rios, Jorge Hallak, José Eduardo Krieger, Juliana C Ferreira, Julio F M Marchini, Larissa S Oliveira, Leila Harima, Linamara R Batisttella, Luis Yu, Luiz Henrique M Castro, Marcelo C Rocha , Marcello M C Magri, Marcio Mancini, Maria Amélia de Jesus, Maria Cassia J M Corrêa, Maria Cristina P B Francisco, Maria Elizabeth Rossi, Marjorie F Silva, Marta Imamura, Maura S Oliveira, Nelson Gouveia, Orestes V Forlenza, Paulo A Lotufo, Ricardo F Bento, Ricardo Nitrini, Rodolfo F Damiano, Roger Chammas, Rossana P Francisco, Solange R G Fusco, Tarcisio E P Barros-Filho, Thais Mauad, Thaís Guimarães, Thiago Avelino-Silva and Wilson J Filho.

Contributors: CRRC: conceptualisation, data curation, formal analysis, funding acquisition, investigation, methodology, project administration, resources, supervision, validation, visualisation, writing – original draft, and writing – review & editing. RCC: data curation, formal analysis, investigation, methodology, validation, visualisation, writing – original draft, and writing – review & editing. MVYS: data curation, formal analysis, investigation, methodology, validation, visualisation, writing – original draft, and writing – review & editing. MLG: conceptualisation, data curation, formal analysis, investigation, methodology, project administration, supervision, validation, visualisation, writing – original draft, and writing – review & editing. CAL: data curation, formal analysis, investigation, writing – original draft, and writing – review & editing. DACC: data curation, formal analysis, methodology, software, writing – original draft, and

writing – review & editing. DML: data curation, formal analysis, methodology, software, writing – original draft, and writing – review & editing. PGS: conceptualisation, project administration, supervision, validation, visualisation and writing – review & editing. JMS: methodology, validation, visualisation and writing – review & editing. CHN: methodology, validation, visualisation and writing – review & editing. MAG: data curation, formal analysis, funding acquisition, methodology, software, supervision, validation, visualisation, writing – original draft, and writing – review & editing. HCFMUSP Covid-19 Study Group: contributed to the implementation of the study and data collection. All authors critically reviewed and approved the final version.

Declaration of interests: We declare no competing interests.

Data Sharing: The study protocol was previously described by Busatto et al. [20] and was registered at the "Brazilian Registry of Clinical Trials" (https://ensaiosclinicos.gov.br/). The raw data are not publicly available because follow-up studies will be carried out. However, data are available from the corresponding author upon request and authorization from the institution. Furthermore, data on demographics, hospitalization, and outcomes are available in the COVID-19 Data Sharing/BR repository and are freely available for download[39].

Funding: Project FAPESP (2020/07200-9) "Analyzing Complex Data Linked to COVID-19 to Support Decision Making and Prognosis".

Competing interests: None declared.

Ethical approval: The study was approved by the Research Ethics Committee of our institution (Process No. 31942020.0.000.0068).

REFERENCES

1. Sisó-Almirall A, Brito-Zerón P, Conangla Ferrín L, et al. Long Covid-19: Proposed Primary Care Clinical Guidelines for Diagnosis and Disease Management. *Int J Environ Res Public Health* 2021;18(8) doi: 10.3390/ijerph18084350 [published Online First: 2021/04/20]
2. Nalbandian A, Sehgal K, Gupta A, et al. Post-acute COVID-19 syndrome. *Nat Med* 2021;27(4):601-15. doi: 10.1038/s41591-021-01283-z [published Online First: 2021/03/22]

3. Mauad T, Duarte-Neto AN, da Silva LFF, et al. Tracking the time course of pathological patterns of lung injury in severe COVID-19. *Respir Res* 2021;22(1):32. doi: 10.1186/s12931-021-01628-9 [published Online First: 20210129]
4. Tanni SE, Fabro AT, de Albuquerque A, et al. Pulmonary fibrosis secondary to COVID-19: a narrative review. *Expert Rev Respir Med* 2021;15(6):791-803. doi: 10.1080/17476348.2021.1916472 [published Online First: 2021/04/27]
5. Macedo BR, Garcia MVF, Garcia ML, et al. Implementation of Tele-ICU during the COVID-19 pandemic. *J Bras Pneumol* 2021;47(2):e20200545. doi: 10.36416/1806-3756/e20200545 [published Online First: 20210430]
6. Huang C, Huang L, Wang Y, et al. 6-month consequences of COVID-19 in patients discharged from hospital: a cohort study. *Lancet* 2021;397(10270):220-32. doi: 10.1016/S0140-6736(20)32656-8 [published Online First: 2021/01/08]
7. Lopez-Leon S, Wegman-Ostrosky T, Perelman C, et al. More Than 50 Long-Term Effects of COVID-19: A Systematic Review and Meta-Analysis. *Res Sq* 2021 doi: 10.21203/rs.3.rs-266574/v1 [published Online First: 2021/03/01]
8. Fernandes PMP, Mariani AW. Life post-COVID-19: symptoms and chronic complications. *Sao Paulo Med J* 2021;139(1):1-2. doi: 10.1590/1516-3180.2021.139104022021
9. Morin L, Savale L, Pham T, et al. Four-Month Clinical Status of a Cohort of Patients After Hospitalization for COVID-19. *JAMA* 2021;325(15):1525-34. doi: 10.1001/jama.2021.3331
10. Raghu G, Collard HR, Egan JJ, et al. An official ATS/ERS/JRS/ALAT statement: idiopathic pulmonary fibrosis: evidence-based guidelines for diagnosis and management. *Am J Respir Crit Care Med* 2011;183(6):788-824. doi: 10.1164/rccm.2009-040GL
11. Carfì A, Bernabei R, Landi F, et al. Persistent Symptoms in Patients After Acute COVID-19. *JAMA* 2020;324(6):603-05. doi: 10.1001/jama.2020.12603
12. Ferreira JC, Ho YL, Besen BAMP, et al. Protective ventilation and outcomes of critically ill patients with COVID-19: a cohort study. *Ann Intensive Care* 2021;11(1):92. doi: 10.1186/s13613-021-00882-w [published Online First: 2021/06/07]
13. WHO. WHO Coronavirus Disease (COVID-19) Dashboard 2021 [cited 2021 October 13]. Available from: https://covid19.who.int/ accessed October 13 2021
14. Wade DT. Rehabilitation after COVID-19: an evidence-based approach. *Clin Med (Lond)* 2020;20(4):359-65. doi: 10.7861/clinmed.2020-0353 [published Online First: 2020/06/09]
15. Godoy CG, Silva ECGE, Oliveira DB, et al. Protocol for Functional Assessment of Adults and Older Adults after Hospitalization for COVID-19. *Clinics (Sao Paulo)* 2021;76:e3030. doi: 10.6061/clinics/2021/e3030 [published Online First: 20210614]
16. Subbaraman N. US health agency will invest $1 billion to investigate 'long COVID'. *Nature* 2021;591(7850):356. doi: 10.1038/d41586-021-00586-y
17. Castiglioni I, Ippolito D, Interlenghi M, et al. Machine learning applied on chest x-ray can aid in the diagnosis of COVID-19: a first experience from Lombardy, Italy. *Eur Radiol Exp* 2021;5(1):7. doi: 10.1186/s41747-020-00203-z [published Online First: 20210202]
18. Wang S, Zha Y, Li W, et al. A fully automatic deep learning system for COVID-19 diagnostic and prognostic analysis. *Eur Respir J* 2020;56(2) doi: 10.1183/13993003.00775-2020 [published Online First: 20200806]
19. Ferreira Junior JR, Cardona Cardenas DA, Moreno RA, et al. Novel Chest Radiographic Biomarkers for COVID-19 Using Radiomic Features Associated with Diagnostics and Outcomes. *J Digit Imaging* 2021;34(2):297-307. doi: 10.1007/s10278-021-00421-w [published Online First: 20210218]
20. Busatto GF, de Araújo AL, Duarte AJDS, et al. Post-acute sequelae of SARS-CoV-2 infection (PASC): a protocol for a multidisciplinary prospective observational evaluation of a cohort of patients surviving hospitalisation in Sao Paulo, Brazil. *BMJ Open*

2021;11(6):e051706. doi: 10.1136/bmjopen-2021-051706 [published Online First: 2021/06/30]
21. van den Borst B, Peters JB, Brink M, et al. Comprehensive health assessment three months after recovery from acute COVID-19. *Clin Infect Dis* 2020 doi: 10.1093/cid/ciaa1750 [published Online First: 2020/11/21]
22. Miller MR, Hankinson J, Brusasco V, et al. Standardisation of spirometry. *Eur Respir J* 2005;26(2):319-38. doi: 10.1183/09031936.05.00034805
23. Pereira CA, Sato T, Rodrigues SC. New reference values for forced spirometry in white adults in Brazil. *J Bras Pneumol* 2007;33(4):397-406. doi: 10.1590/s1806-37132007000400008
24. Hansell DM, Bankier AA, MacMahon H, et al. Fleischner Society: glossary of terms for thoracic imaging. *Radiology* 2008;246(3):697-722. doi: 10.1148/radiol.2462070712 [published Online First: 20080114]
25. Francone M, Iafrate F, Masci GM, et al. Chest CT score in COVID-19 patients: correlation with disease severity and short-term prognosis. *Eur Radiol* 2020;30(12):6808-17. doi: 10.1007/s00330-020-07033-y [published Online First: 2020/07/04]
26. Wu Q, Zhong L, Li H, et al. A Follow-Up Study of Lung Function and Chest Computed Tomography at 6 Months after Discharge in Patients with Coronavirus Disease 2019. *Can Respir J* 2021;2021:6692409. doi: 10.1155/2021/6692409 [published Online First: 20210213]
27. Huang L, Yao Q, Gu X, et al. 1-year outcomes in hospital survivors with COVID-19: a longitudinal cohort study. *Lancet* 2021;398(10302):747-58. doi: 10.1016/S0140-6736(21)01755-4
28. Han X, Fan Y, Alwalid O, et al. Fibrotic Interstitial Lung Abnormalities at 1-year Follow-up CT after Severe COVID-19. *Radiology* 2021:210972. doi: 10.1148/radiol.2021210972 [published Online First: 20210727]
29. Han X, Fan Y, Alwalid O, et al. Six-month Follow-up Chest CT Findings after Severe COVID-19 Pneumonia. *Radiology* 2021;299(1):E177-E86. doi: 10.1148/radiol.2021203153 [published Online First: 20210126]
30. Stylemans D, Smet J, Hanon S, et al. Evolution of lung function and chest CT 6 months after COVID-19 pneumonia: Real-life data from a Belgian University Hospital. *Respir Med* 2021;182:106421. doi: 10.1016/j.rmed.2021.106421 [published Online First: 20210418]
31. Caruso D, Guido G, Zerunian M, et al. Post-Acute Sequelae of COVID-19 Pneumonia: Six-month Chest CT Follow-up. *Radiology* 2021;301(2):E396-E405. doi: 10.1148/radiol.2021210834 [published Online First: 20210727]
32. Guler SA, Ebner L, Aubry-Beigelman C, et al. Pulmonary function and radiological features 4 months after COVID-19: first results from the national prospective observational Swiss COVID-19 lung study. *Eur Respir J* 2021;57(4) doi: 10.1183/13993003.03690-2020 [published Online First: 2021/04/29]
33. Liu C, Ye L, Xia R, et al. Chest Computed Tomography and Clinical Follow-Up of Discharged Patients with COVID-19 in Wenzhou City, Zhejiang, China. *Ann Am Thorac Soc* 2020;17(10):1231-37. doi: 10.1513/AnnalsATS.202004-324OC
34. Tabatabaei SMH, Rajebi H, Moghaddas F, et al. Chest CT in COVID-19 pneumonia: what are the findings in mid-term follow-up? *Emerg Radiol* 2020;27(6):711-19. doi: 10.1007/s10140-020-01869-z [published Online First: 20201109]
35. Solomon JJ, Heyman B, Ko JP, et al. CT of Postacute Lung Complications of COVID-19. *Radiology* 2021:211396. doi: 10.1148/radiol.2021211396 [published Online First: 20210810]
36. Sonnweber T, Sahanic S, Pizzini A, et al. Cardiopulmonary recovery after COVID-19: an observational prospective multicentre trial. *Eur Respir J* 2021;57(4) doi: 10.1183/13993003.03481-2020 [published Online First: 2021/04/29]
37. DataSUS. MdS. SIGTAP - Sistema de Gerenciamento da Tabela de Procedimentos, Medicamentos e OPM do SUS. 2021 [cited 2021 January 03]. Available from:

http://sigtap.datasus.gov.br/tabela-unificada/app/sec/inicio.jsp accessed January 03 2021.

38. Mettler FAM, Mahadevappa Bhargavan, MythreyiChambers, Charles E.Elee, Jennifer G. Frush, Donald P. Milano, Michael T. Miller, Donald L. Royal, Henry D. Spelic, David C.Ansari, Armin J. Bolch, Wesley E.Guebert, Gary M. Sherrier, Robert H. Smith, James M. Vetter, Richard J. Report Na. 184 - Medical Radiation Exposure of Patients in the United States. United States: National Council on Radiation Protection and Measurements, 2019.
39. FAPESP. COVID-19 DataSharing/BR 2021 [cited 2021 September 08]. Available from: https://repositoriodatasharingfapesp.uspdigital.usp.br/.

Figure Legends

Figure 1. Logistic regression-based machine learning model. The modified Medical Research Council (mMRC) dyspnoea scale, oximetry (SpO2), spirometry (forced vital capacity [FVC]), and the five radiographic scores obtained during DL-based classification of CXR (pRX) were used as input data, and the presence of lung lesions due to COVID-19 was used as output data. AI: artificial intelligence. CT: computed tomography.

Figure 2. Flowchart of patient selection. FVC: forced vital capacity; LLN; lower limit of normal; mMRC, modified Medical Research Council dyspnoea scale. *Rest SpO_2 < 90% or a decrease in SpO_2 of at least 4% after the 1-min sit and stand test.

Figure 3. Fibrotic-like changes after a critical COVID-19 of a patient in his early 70s. (A) PA chest radiograph obtained 7 months after infection shows reticular opacities with a slight peripheral predominance diffusely distributed in both lungs. (B) Image from the same radiograph analysed by the AI algorithm with heat map highlighting the areas of pulmonary involvement. (C, D) Chest CT obtained 8 months after infection shows moderate ground glass opacities, linear multifocal and reticular abnormalities, discrete traction bronchiectasis and slight parenchymal architectural distortion. The patient had dyspnoea (mMRC=1) and altered FVC (2.34 L / 60% pred), besides the normal oximetry (97%).

Tables

Table 1. Pulmonary function of patients with signs of pulmonary involvement.[a]	
Variables	**Patients with signs of pulmonary involvement (N=749)**
mMRC ≥ 2	229/742 (30.9%)
Altered Oximetry*	71/675 (10.5%)
CRX (score 1)	200/629 (31.8%)
FVC < LLN	212/642 (33%)

CRX: chest X-ray; FVC: forced vital capacity; mMRC: modified Medical Research Council dyspnoea scale. LLN, lower limit of normal. [a]Values are n/N (%). *Resting SpO_2 ≤90% or a decrease in SpO_2 of ≥4% during the 1-min sit and stand test.

Table 2. Performance of the predictive model using three combinations of variables.[a]

Groups of variables	Sensitivity	Specificity	F1-score	AUC
1 SpO_2, mMRC score, and FVC	0.87±0.16	0.42±0.33	0.71±0.03	0.68±0.10
2 CRX	0.88±0.05	0.52±0.14	0.75±0.04	0.78±0.05
3 SpO_2, mMRC score, FVC, and CRX	0.85±0.08	0.70±0.14	0.79±0.06	0.80±0.07

CRX: chest X-Ray; mMRC: Modified Medical Research Council dyspnoea scale; FVC: forced vital capacity. [a]values are means ± standard deviations after 5-fold cross validation for each test fold.

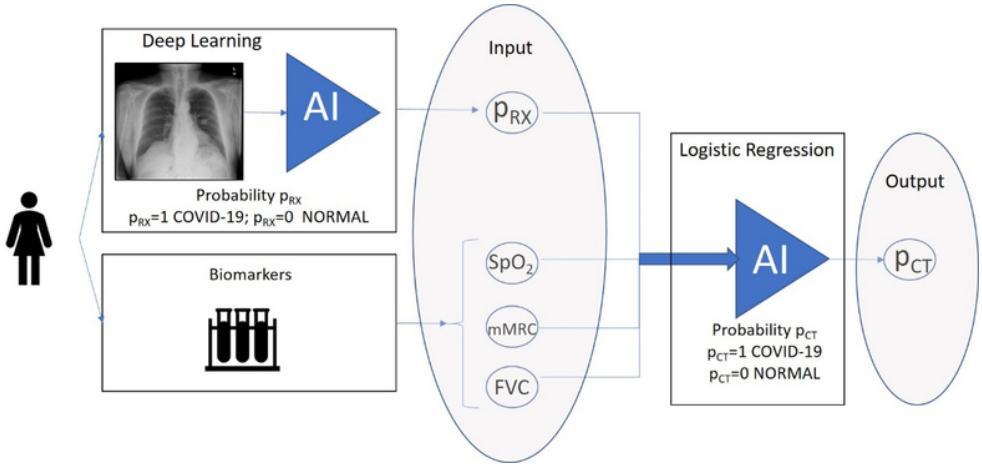


Figure 1. Logistic regression-based machine learning model. The modified Medical Research Council (mMRC) dyspnoea scale, oximetry (SpO2), spirometry (forced vital capacity [FVC]), and the five radiographic scores obtained during DL-based classification of CXR (pRX) were used as input data, and the presence of lung lesions due to COVID-19 was used as output data. AI: artificial intelligence. CT: computed tomography.

67x31mm (300 x 300 DPI)

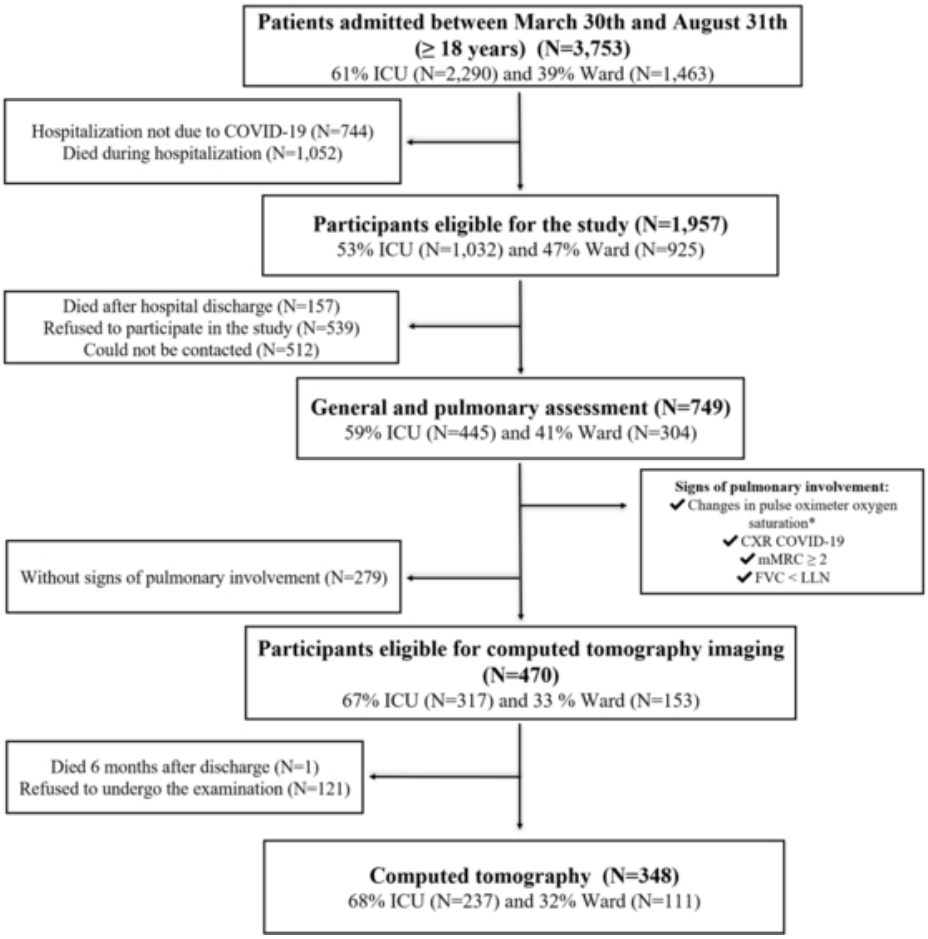


Figure 2. Flowchart of patient selection. FVC: forced vital capacity; LLN; lower limit of normal; mMRC, modified Medical Research Council dyspnoea scale. *Rest SpO2 < 90% or a decrease in SpO2 of at least 4% after the 1-min sit and stand test.

50x47mm (300 x 300 DPI)

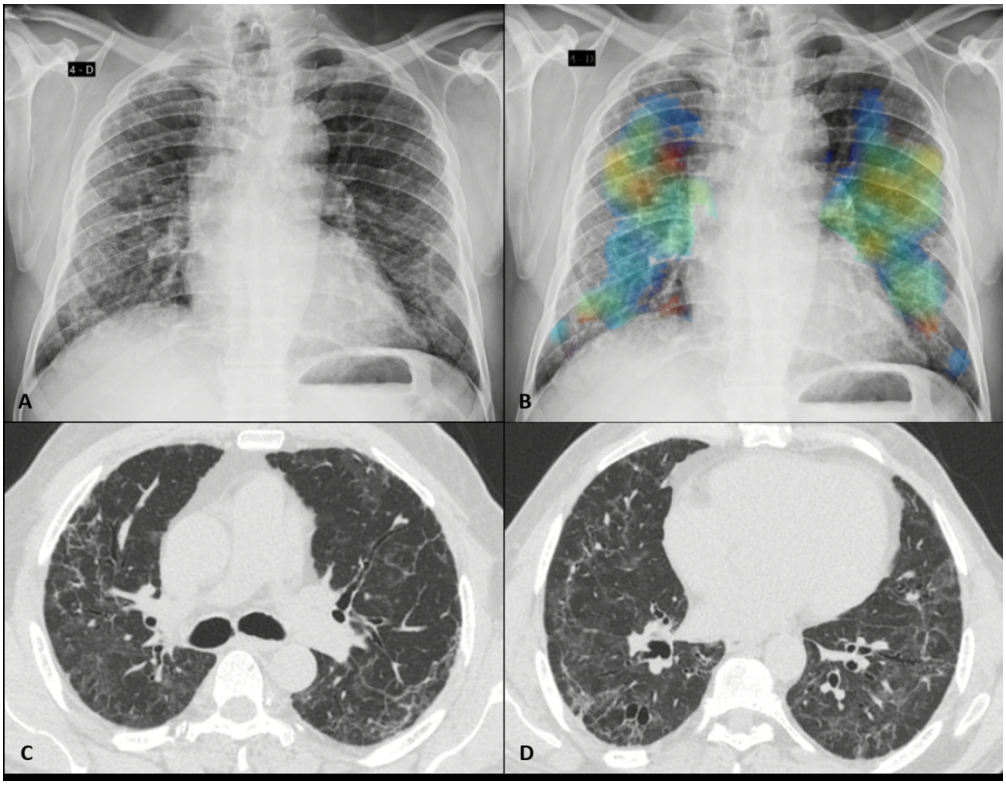


Figure 3. Fibrotic-like changes after a critical COVID-19 of a patient in his early 70s. (A) PA chest radiograph obtained 7 months after infection shows reticular opacities with a slight peripheral predominance diffusely distributed in both lungs. (B) Image from the same radiograph analysed by the AI algorithm with heat map highlighting the areas of pulmonary involvement. (C, D) Chest CT obtained 8 months after infection shows moderate ground glass opacities, linear multifocal and reticular abnormalities, discrete traction bronchiectasis and slight parenchymal architectural distortion. The patient had dyspnoea (mMRC=1) and altered FVC (2.34 L / 60% pred), besides the normal oximetry (97%).

154x119mm (300 x 300 DPI)

Data Supplement

Chronic lung lesions in COVID-19 survivors: predictive clinical model

Carlos R R Carvalho, Rodrigo C Chate, Marcio VY Sawamura, Michelle L Garcia, Celina A Lamas, Diego AC Cardenas, Daniel M Lima, Paula G Scudeller, João M Salge, Cesar H Nomura, Marco A Gutierrez, HCFMUSP Covid-19 Study Group.

Contents

Supplementary Methods

Datasets

The SIIM-RSNA dataset contains 6,334 posterior-anterior radiographic images from 6,054 patients obtained from the public dataset Machine Learning Challenge on COVID-19 Pneumonia Detection and Localization [1]. Specialists classified images as “negative for pneumonia” or “COVID-19 pneumonia”. A total of 6,030 images were selected and randomly distributed in training and validation sets (1,276 negative and 3,711 positive) and in a test set (400 negative and 643 positive).

The InRad dataset contains chest X-Ray (CXR) and chest computed tomographic (CT) images of 257 patients. The CXR images were classified as normal (145 patients) or with findings related to COVID-19 (112 patients) and randomly distributed in training and validation sets (214 patients) and a test set (43 patients). Images were obtained from the Institute of Radiology (InRad) of the Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo (HCFMUSP).

Because of differences in dataset sizes, a data augmentation technique was adopted using random transformations, including rotation (0–15 degrees), horizontal mirroring, and random changes in intensity and contrast (0–5%).

Classification of chest radiography images

A Deep Learning (DL) approach using a Convolutional Neural Network (CNN) based on an EfficientNetB7 architecture was used [2]. The network classification layer was replaced by a Global Average Pooling operation, followed by Batch Normalization and the adoption of a dense layer with one neuron and sigmoid activation function. Each training iteration was run for 40 epochs with an Adam optimizer at a learning rate of 0.0001. All images were resized to 600 ´ 600 pixels.

The CNN was trained using the SIIM-RSNA dataset to detect radiographic patterns of COVID-19 pneumonia. The training was initiated in EfficientNetB7 using weights after pre-training with the ImageNet dataset [3].

A 5-fold cross-validation strategy was adopted over the training and validation sets. The training weights obtained for each fold were used with the test set of the SIIM-SNA to describe the classification accuracy (Table 1). The fold with the best result in terms of the metric the area under the receiver operating characteristic curve (AUC), in this case, fold 1 with AUC of 0.89, defines the final weights of the CNN.

Table 1. Classification of the test set of the SIIM-RSNA dataset as negative (normal) or positive (patterns of COVID-19 pneumonia); Accuracy (Acc); Precision (Prec).

Dataset	5-fold	Acc	Prec	Sensitivity	Specificity	F1-score	AUC
SIIM-RSNA	0	0,80	0,85	0,82	0,76	0,83	0,88
	1	0,80	0,85	0,82	0,77	0,84	**0,89**
	2	0,78	0,77	0,92	0,56	0,84	0,87
	3	0,76	0,74	0,93	0,48	0,83	0,86
	4	0,76	0,74	0,93	0,48	0,83	0,86

For the InRad dataset, the CNN was initialized with the final weights defined in training with SIIM-RSNA. After initialization, the CNN was retrained to classify images as normal or with finds related to COVID-19.

The InRad dataset was divided into six-folds during the retraining, five folds for training and validation, and one-fold for test. In order to avoid bias, the test fold was selected to run all six folds available and, for each test fold selected, a 5-fold cross-validation strategy was applied in the remaining training and validation folds (Table 2).

Table 2. Classification using six test folds of the InRad database. For each test fold, the values for the metrics Accuracy (Acc), Precision (Prec), Sensitivity, Specificity, F1-score and AUC represent the mean and standard deviation after 5-fold cross validation

Dataset	Test fold	Acc	Prec	Sensitivity	Specificity	F1-score	AUC
InRad	0	0.79±0.01	0.74±0.04	0.82±0.07	0.77±0.06	0.78±0.02	0.86±0.02
	1	0.69±0.02	0.62±0.03	0.84±0.06	0.57±0.07	0.71±0.02	0.75±0.01
	2	0.67±0.05	0.60±0.06	0.81±0.08	0.57±0.13	0.68±0.02	0.76±0.02
	3	0.77±0.04	0.71±0.07	0.80±0.04	0.74±0.10	0.75±0.03	0.80±0.02
	4	0.82±0.05	0.77±0.11	0.89±0.10	0.78±0.14	0.81±0.03	0.89±0.04
	5	0.71±0.04	0.62±0.04	0.90±0.02	0.58±0.08	0.73±0.03	0.80±0.02

Detection of chronic lung lesions on computed tomography images

Three machine learning models were developed using as input the clinical data (modified Medical Research Council dyspnea scale [mMRC], oximetry [SpO_2] and spirometry [forced vital capacity, FVC]), and five radiographic probabilities (p_{RX0} to p_{RX4}) with findings related to COVID-19 ($p_{RXn}=1$) and normal ($p_{RXn}=0$), obtained from the previous step (Table 2). As output, the models predict the value of a binary variable (p_{CT}) related to the presence of chronic lung lesions on CT images, with $p_{CT}=1$ for a CT score ≥ 7 (129 patients) and $p_{CT}=0$ for a CT score < 7 (128 patients) (Figure 1).

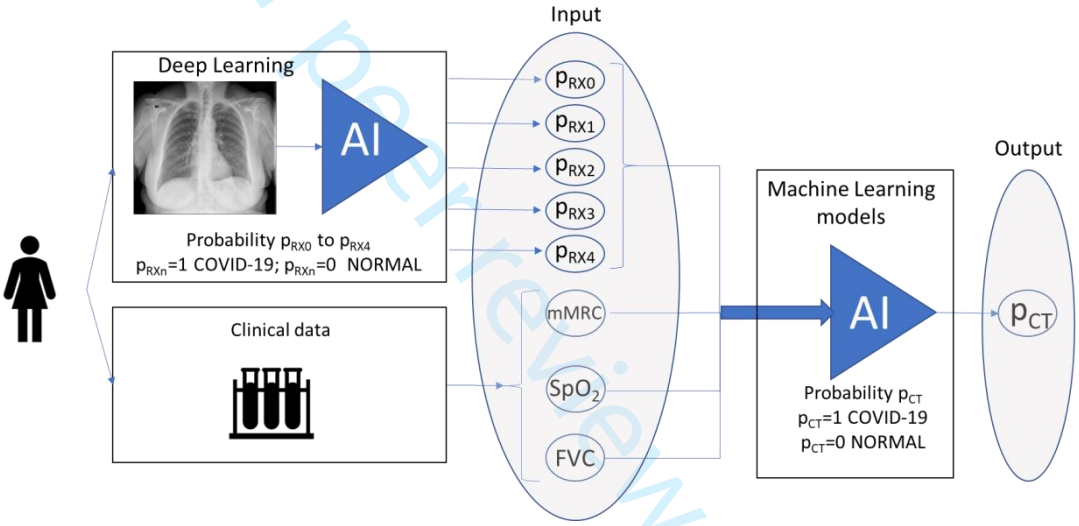


Figure 1. Machine learning-based model. The modified Medical Research Council (mMRC) dyspnea scale, oximetry (SpO2), spirometry (forced vital capacity [FVC]), and the radiographic probabilities (p_{RX0} to p_{RX4}) with findings related to COVID-19 (pRXn=1) and normal (pRXn=0) as input data, and the presence of lung lesions due to COVID-19 (p_{CT}) was used as output data; AI: artificial intelligence. CT: computed tomography.

The first model was LogisticRegression (LR) with L2 regularization (4). The second model was RandomForest with 100 trees (RF-100), Gini criterion, minimum of two samples for splitting, minimum of one sample in leaves, and bootstrap (4). The third model was RandomForest with the parameters described above, except for the limit of 10 trees and maximum depth h_max=6 (RF-10) [4]. The performance of the machine learning models was evaluated by the metrics Sensitivity, Specificity, AUC, and F1-score.

Three combinations of input variables were evaluated: 1) clinical variables (mMRC, SpO_2, and FVC); 2) CXR; 3) clinical variables (mMRC, SpO_2, FVC and CXR.

The performance of the logistic regression (LR) model was higher using the combination of all variables (clinical variables and CXR) considering the metrics Sensitivity of 0.85±0.08 (95% CI [0.77,0.94]), Specificity of 0.70±0.14 (95% CI [0.55, 0.85]), F1-score of 0.79±0.06 (95% CI [0.73, 0.85]), and AUC of 0.80±0.07(95% CI [0.72, 0.87]), expressed in terms of mean and standard deviation (Table 3).

Using the LR model, the predictive model is represented by the following function:

$$p_{CT} = 0.59 \times \left(\frac{FVC_{Resting}}{2FVC_{lin}}\right) - 2.16 \times \left(\frac{mMRC}{4}\right) + 0{,}679(SpO_2) + 1.15 \times p_{RX0} + 1.41 \times p_{RX1} + 1.04 \times p_{RX2} + 0.69 \times p_{RX3} + 0.60 \times p_{RX4}$$

where p_{CT} is the presence of abnormalities on CT images.

Table 3. Predictive performance of three multivariate models using three datasets.[a]

Groups of variables	Method	Sensitivity	Specificity	F1-score	AUC
1 SpO_2, mMRC score, and FVC	LR	0.87±0.16	0.42±0.33	0.71±0.03	0.68±0.10
	RF-10	0.88±0.15	0.37±0.32	0.71±0.03	0.66±0.08
	RF-100	0.82±0.12	0.44±0.13	0.69±0.08	0.62±0.12
2 CXR	LR	0.88±0.05	0.52±0.14	0.75±0.04	0.78±0.05
	RF-10	0.91±0.08	0.41±0.18	0.73±0.04	0.73±0.06
	RF-100	0.94±0.07	0.33±0.19	0.72±0.03	0.72±0.03
3 SpO_2, mMRC score, FVC and CRX	LR	**0.85±0.08**	**0.70±0.14**	**0.79±0.06**	**0.80±0.07**
	RF-10	0.85±0.09	0.61±0.22	0.76±0.04	0.76±0.08
	RF-100	0.89±0.06	0.49±0.17	0.75±0.04	0.76±0.07

CRX: chest X-Ray; mMRC: modified Medical Research Council dyspnea scale (mMRC); FVC: forced vital capacity. [a]values are means ± standard deviations after 5-fold cross validation for each test fold.

Dataset and normalization of clinical data

A total of 257 patients with data for the mMRC dyspnea scale, oximetry, spirometry, CRX, and chest CT were selected to predict pulmonary changes. Of the 257 patients, 128 had no significant CT changes (scores < 7). A CT score of 7 was used as the cutoff value by maximizing F1 scores and AUC (Figure 2).

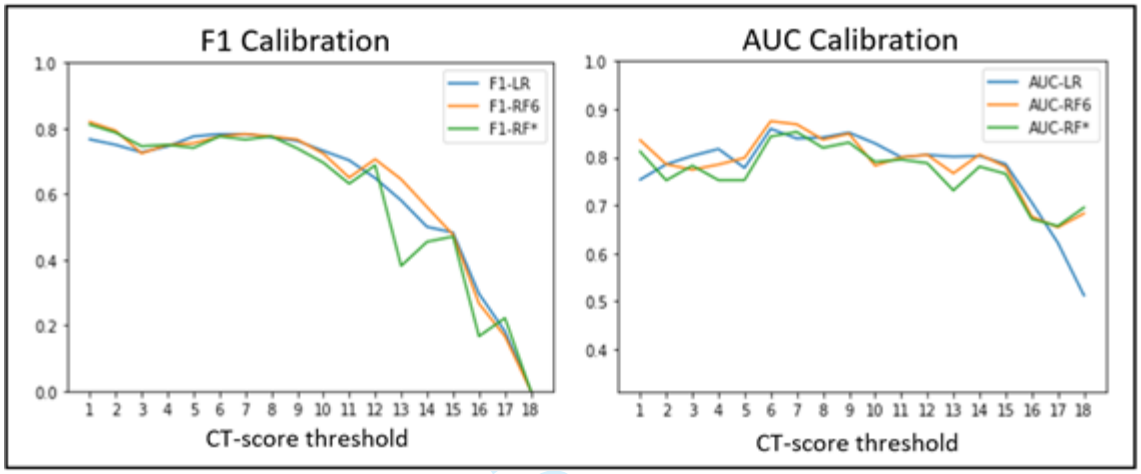
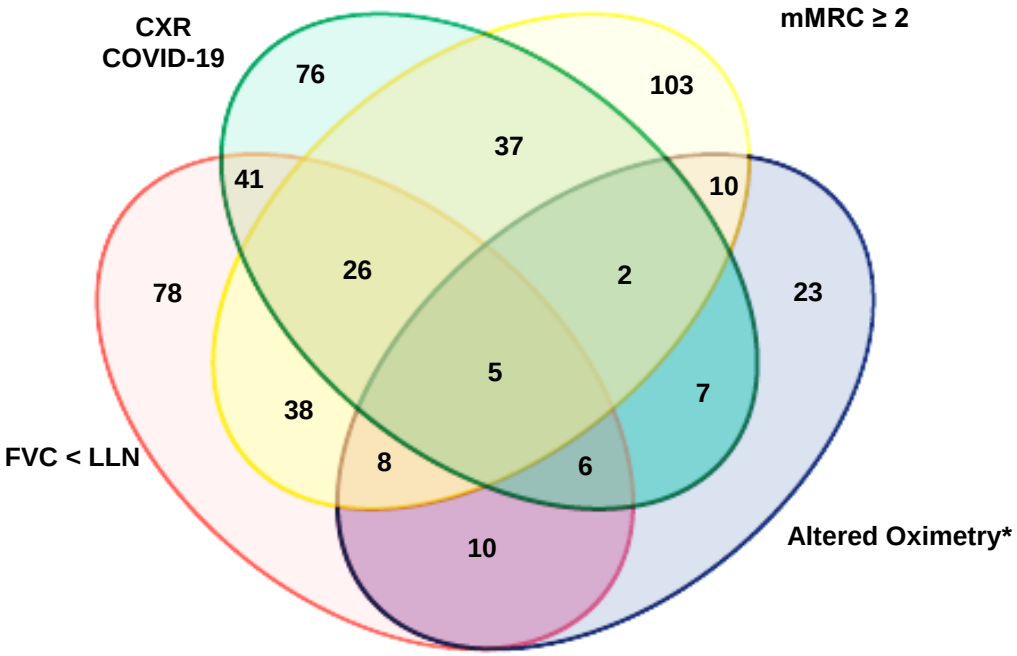
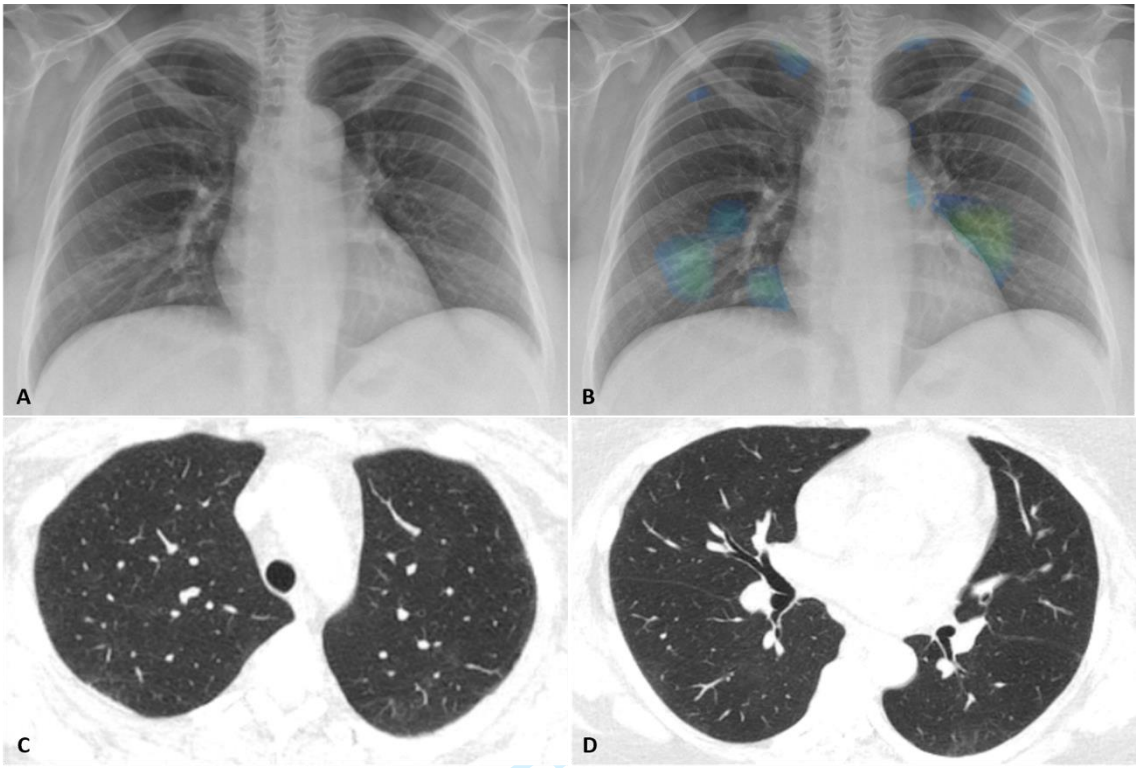


Figure 2. Computed tomography scores based on F1-score and AUC values.

Clinical variables were normalized by dividing mMRC values by 4 (resulting in values between 0 and 1) and the $FVC_{Resting}$ by twice the FVC_{lln} (resulting in a minimum value of 0.257 and a maximum value of 0.847).



Supplementary Figure S1. Signs of pulmonary involvement. Values are expressed as N. CXR: chest X-Ray. FVC: forced vital capacity. LLN: lower limit of normal. mMRC: modified Medical Research Council dyspnea scale. *Resting SpO_2 ≤ 90% or a decrease in SpO_2 of ≥ 4% during the 1 min sit-and-stand test.



Supplementary Figure 2. Resolving ground glass abnormality after moderate COVID-19 of a patient in her late 40s. (A) PA chest radiograph obtained 8 months after admission was considered normal in the analysis of the radiologists. (B) Image from the same radiograph analyzed by the AI algorithm with heat map highlighting small focal abnormalities in the apical and paracardiac regions of the lungs. (C, D) Chest CT obtained 11 months after admission showing mild residual ground glass abnormality in the periphery of the upper lobes and left lower lobe. The patient complained of dyspnea (mMRC=3) but had normal lung function (FVC = 3.81 L / 91% pred) and normal oximetry (99%).

Supplementary Table S1. Demographic and clinical data of a population of post-COVID-19 patients.[a]

Variables	N=749
Age (years)	56.1 (44.4–65.1)
Male sex	399 (53.3%)
BMI (kg/m^2)	30.8 (27.7–35.6) {746}
Comorbidities	
Hypertension	425 (56.7%)
Smokers	285/743 (38.4%)
Diabetes	261 (34.8%)
COPD	55 (7.3%)
Admission	
ICU	445 (59.4%)
Length of ICU stay (days)	10 (6–18) {445}
IMV	304/445 (68.3%)
Vital signs	
Body temperature (°C)	36.1 (35.6–36.0) {748}
Systolic blood pressure (mmHg)	124 (116–135) {743}
Diastolic blood pressure (mmHg)	77 (70–84) {743}
Heart rate (bpm)	73 (67–83) {747}
Respiratory rate (rpm)	20 (18–2) {736}
Oxygen saturation (%)	97 (95.2–98) {746}

COPD: chronic obstructive pulmonary disease; BMI: body mass index; ICU: intensive care unit. IMV: invasive mechanical ventilation. [a]Values are median (IQR), median (IQR) {n}, n (%), or n/N (%).

Supplementary Table S2. Demographic and clinical characteristics of patients with and without pulmonary involvement.[a]

Variables	Pulmonary involvement (n=470)	No pulmonary involvement (n=279)	p-value
Age (years)	57.9 (45.7–65.8)	53.9 (42.5–63.7)	0.0005
Male sex	228 (48.5%)	171 (61.3%)	0.0007
BMI (kg/m^2)	31.2 (27.7–35.9) {469}	30.5 (27.6–35.2) {277}	0.1112
Comorbidities			
Hypertension	287 (61.1%)	138 (49.5%)	0.0005
Smokers	188/468 (40.2%)	97/275 (35.3%)	0.1039
Diabetes	179 (38.1%)	82 (29.4%)	0.0092
COPD	42 (8.9%)	13 (4.7%)	0.0445
Admission			
ICU	317 (67.4%)	128 (45.9%)	0.0001
Length of ICU stay (days)	11 (6–20) {317}	8 (4–14) {128}	0.0001
IMV	222/317 (70%)	82/128 (64.1%)	0.2603

COPD: chronic obstructive pulmonary disease; BMI: body mass index; ICU: intensive care unit. IMV: invasive mechanical ventilation.
[a]Values are median (IQR), median (IQR) {n}, n (%), or n/N (%).

Supplementary Table S3. Demographic and clinical data of COVID-19 patients with signs of pulmonary involvement that underwent or did not undergo the chest computed tomography exam.[a]

Variables	Presence of pulmonary involvement		p-value
	Underwent CT (n=348)	Did not undergo CT (n=122)	
Age (years)	57.8 (45.7–65.8)	58.1 (45.3–65.8)	0.49
Male sex	163 (46.8%)	65 (53.3%)	0.3922
BMI (kg/m^2)	31.6 (28.0–36.0)	30.3 (27.0–35.9) {121}	0.0407
Comorbidities			
Hypertension	215 (61.8%)	72 (59%)	0.4691
Smokers	139/347 (40.1%)	49/121 (40.5%)	0.7619
Diabetes	142 (40.8%)	37 (30.3%)	0.9999
COPD	32 (9.2%)	10 (8.2%)	0.826
Admission			
ICU	237 (68.1%)	80 (65.6%)	0.9999
Length of ICU stay (days)	11 (6–20) {237}	10 (4.7–19) {80}	0.9133
IMV	174/237 (73.4%)	48/80 (60%)	0.0337

COPD: chronic obstructive pulmonary disease; BMI: body mass index; ICU: intensive care unit. IMV: invasive mechanical ventilation.
[a]Values are median (IQR), median (IQR) {n}, n (%), or n/N (%).

Supplementary Table S4. Chest computed tomography (CT) features in a population of COVID-19 patients.	
Variables	**CT changes**
CT score ≥ 7	156/328 (47.6%)
Characteristics (n=156)	
Ground-glass opacities	153 (98.1%)
Parenchymal bands	143 (91.7%)
Reticulations	134 (85.9%)
Traction bronchiectasis	92 (59.0%)
Architectural distortion	73 (46.8%)
Perilobular opacities	50 (32.1%)
Bronchial wall thickening	38 (24.4%)
Mosaic attenuation pattern	32 (20.5%)
Consolidations	3 (1.9%)
Pneumatocele	2 (1.3%)
Honeycombing	-
[a]Values are n/N (%) or n (%).	

Supplementary Table S5. Computed tomography changes 6 to 11 months after hospitalization due to COVID-19.[a]			
Characteristics	**Total cohort (N=328)**	**ICU Patients (N=222)**	**Ward Patients (N=106)**
Ground-glass opacities	251 (76.5%)	197 (86.6%)	54 (51.3%)
Parenchymal bands	209 (63.7%)	169 (76.5%)	40 (41%)
Reticulations	169 (51.5%)	145 (66.5%)	24 (23.1%)
Traction bronchiectasis	98 (29.9%)	91 (44.1%)	7 (7.7%)
Architectural distortion	78 (23.8%)	73 (35.8%)	5 (6.4%)
Bronchial wall thickening	89 (27.1%)	60 (27.4%)	29 (25.6%)
Mosaic attenuation pattern	58 (17.7%)	46 (20.1%)	12 (11.5%)
Perilobular opacities	50 (14%)	47 (24.6%)	3 (2.6%)
Consolidation	3 (0.9%)	3 (1.7%)	-
Pneumatocele	2 (0.6%)	2 (1.1%)	-
Honeycombing	-	-	-
[a]Values are n (%)			

Supplementary Table S6. Demographic and clinical data of COVID-19 patients with pulmonary involvement included or excluded from the analysis of prediction of pulmonary changes.[a]

Variables	Prediction of Pulmonary Changes		p-value
	Patients Included (N=257)	Patients Excluded (N=91)	
Age (years)	56.5 (45.7–64.4)	60.5 (46.9–69.9)	0.011
Male sex	113 (44%)	50 (54.9%)	0.0681
BMI (kg/m^2)	32 (28.8–36.8)	30.6 (26.8–35.4)	0.0537
Comorbidities			
Hypertension	151 (58.7%)	64 (70.3%)	0.0601
Smokers	97/256 (37.9%)	42 (46.1%)	0.173
Diabetes	103 (40.1%)	39 (42.9%)	0.7101
COPD	20 (7.8%)	12 (13.2%)	0.1415
Admission			
ICU	179 (69.6%)	58 (63.7%)	0.3598
Length of ICU stay (days)	12 (6–20.5) {179}	9.5 (6.2–19.7) {58}	0.209
IMV	140 (54.7%)	35 (38.6%)	0.0105

BMI: body mass index; COPD: chronic obstructive pulmonary disease; ICU: intensive care unit. IMV: invasive mechanical ventilation.
[a]Values are median (IQR), median (IQR) {n}, n (%), or n/N (%).

Supplementary References

1. Stephens K. SIIM, FISABIO, and RSNA Host Machine Learning Challenge for COVID-19 Detection and Localization. . *AXIS Imaging News* . 2021.

2. Tan M, Le Q. Efficientnet: Rethinking model scaling for convolutional neural networks. *International Conference on Machine Learning* 2019. p. 6105-14.

3. Russakovsky O, Deng J, Su H, et al. ImageNet Large Scale Visual Recognition Challenge. *International Journal of Computer Vision* 2015; **115**: 211-52.

4. James, G., Witten D, Hastie T, Tibshirani R. *An introduction to statistical learning* New York: Springer; 2013.

TRIPOD Checklist: Prediction Model Development and Validation

Section/Topic	Item		Checklist Item	Page
Title and abstract				
Title	1	D;V	Identify the study as developing and/or validating a multivariable prediction model, the target population, and the outcome to be predicted.	1
Abstract	2	D;V	Provide a summary of objectives, study design, setting, participants, sample size, predictors, outcome, statistical analysis, results, and conclusions.	2
Introduction				
Background and objectives	3a	D;V	Explain the medical context (including whether diagnostic or prognostic) and rationale for developing or validating the multivariable prediction model, including references to existing models.	4
	3b	D;V	Specify the objectives, including whether the study describes the development or validation of the model or both.	5
Methods				
Source of data	4a	D;V	Describe the study design or source of data (e.g., randomized trial, cohort, or registry data), separately for the development and validation data sets, if applicable.	5
	4b	D;V	Specify the key study dates, including start of accrual; end of accrual; and, if applicable, end of follow-up.	5
Participants	5a	D;V	Specify key elements of the study setting (e.g., primary care, secondary care, general population) including number and location of centres.	5
	5b	D;V	Describe eligibility criteria for participants.	5
	5c	D;V	Give details of treatments received, if relevant.	-
Outcome	6a	D;V	Clearly define the outcome that is predicted by the prediction model, including how and when assessed.	7
	6b	D;V	Report any actions to blind assessment of the outcome to be predicted.	Data Supplement (Pg. 3 and 5)
Predictors	7a	D;V	Clearly define all predictors used in developing or validating the multivariable prediction model, including how and when they were measured.	6, 7
	7b	D;V	Report any actions to blind assessment of predictors for the outcome and other predictors.	Data Supplement (Pg. 3 and 5)
Sample size	8	D;V	Explain how the study size was arrived at.	5
Missing data	9	D;V	Describe how missing data were handled (e.g., complete-case analysis, single imputation, multiple imputation) with details of any imputation method.	Data Supplement, (Pg. 6)
Statistical analysis methods	10a	D	Describe how predictors were handled in the analyses.	Data Supplement (Pg. 3 and 5)
	10b	D	Specify type of model, all model-building procedures (including any predictor selection), and method for internal validation.	Data Supplement (Pg. 3 and 5)
	10c	V	For validation, describe how the predictions were calculated.	Data Supplement (Pg. 3 and 5)
	10d	D;V	Specify all measures used to assess model performance and, if relevant, to compare multiple models.	Data Supplement (Pg. 3 and 5)
	10e	V	Describe any model updating (e.g., recalibration) arising from the validation, if done.	S Data Supplement (Pg. 4)
Risk groups	11	D;V	Provide details on how risk groups were created, if done.	n.a.
Development vs. validation	12	V	For validation, identify any differences from the development data in setting, eligibility criteria, outcome, and predictors.	Data Supplement (Pg. 3 and 5)
Results				
Participants	13a	D;V	Describe the flow of participants through the study, including the number of participants with and without the outcome and, if applicable, a summary of the follow-up time. A diagram may be helpful.	8
	13b	D;V	Describe the characteristics of the participants (basic demographics, clinical features, available predictors), including the number of participants with missing data for predictors and outcome.	8
	13c	V	For validation, show a comparison with the development data of the distribution of important variables (demographics, predictors and outcome).	Data Supplement (Pg. 3 and 5)
Model development	14a	D	Specify the number of participants and outcome events in each analysis.	Data Supplement (Pg. 3 and 5)
	14b	D	If done, report the unadjusted association between each candidate predictor and outcome.	Data Supplement (Pg. 3 and 5)
Model specification	15a	D	Present the full prediction model to allow predictions for individuals (i.e., all regression coefficients, and model intercept or baseline survival at a given time point).	Data Supplement (Pg. 3 and 5)

TRIPOD Checklist: Prediction Model Development and Validation

	15b	D	Explain how to the use the prediction model.	Data Supplement (Pg. 3 and 5)
Model performance	16	D;V	Report performance measures (with CIs) for the prediction model.	Data Supplement (Pg. 5)
Model-updating	17	V	If done, report the results from any model updating (i.e., model specification, model performance).	Data Supplement (Pg.5)
Discussion				
Limitations	18	D;V	Discuss any limitations of the study (such as nonrepresentative sample, few events per predictor, missing data).	12
Interpretation	19a	V	For validation, discuss the results with reference to performance in the development data, and any other validation data.	Data Supplement (Pg.5)
	19b	D;V	Give an overall interpretation of the results, considering objectives, limitations, results from similar studies, and other relevant evidence.	10, 11, 12
Implications	20	D;V	Discuss the potential clinical use of the model and implications for future research.	10, 11, 12
Other information				
Supplementary information	21	D;V	Provide information about the availability of supplementary resources, such as study protocol, Web calculator, and data sets.	n.a.
Funding	22	D;V	Give the source of funding and the role of the funders for the present study.	14

*Items relevant only to the development of a prediction model are denoted by D, items relating solely to a validation of a prediction model are denoted by V, and items relating to both are denoted D;V. We recommend using the TRIPOD Checklist in conjunction with the TRIPOD Explanation and Elaboration document.

BMJ Open

Chronic lung lesions in COVID-19 survivors: predictive clinical model

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2021-059110.R1
Article Type:	Original research
Date Submitted by the Author:	23-Dec-2021
Complete List of Authors:	Carvalho, Carlos; Universidade de São Paulo, Instituto do Coração - Divisão de Pneumologia Chate, Rodrigo; Universidade de São Paulo Hospital das Clínicas, Instituto de Radiologia Sawamura, Marcio; Universidade de Sao Paulo Hospital das Clinicas, Instituto de Radiologia Garcia, Michelle; Universidade de São Paulo Hospital das Clínicas, Instituto do Coração - Divisão de Pneumologia Lamas, Celina ; Universidade de Sao Paulo Hospital das Clinicas, Instituto do Coração - Divisão de Pneumologia Cardenas, Diego; Universidade de São Paulo Hospital das Clínicas, Instituto do Coração - Divisão de Informática Lima, Daniel Mario; Universidade de São Paulo Hospital das Clínicas, Instituto do Coração - Divisão de Informática Scudeller, Paula; Universidade de São Paulo Hospital das Clínicas, Instituto do Coração - Divisão de Pneumologia Salge, João; Universidade de São Paulo Hospital das Clínicas, Instituto do Coração - Divisão de Pneumologia Nomura, Cesar; Universidade de São Paulo Hospital das Clínicas, Instituto de Radiologia Gutierrez, Marco; Universidade de São Paulo Hospital das Clínicas, Instituto do Coração - Divisão de Pneumologia
Primary Subject Heading:	Respiratory medicine
Secondary Subject Heading:	Global health
Keywords:	COVID-19, Chest imaging < RADIOLOGY & IMAGING, RESPIRATORY MEDICINE (see Thoracic Medicine)

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Chronic lung lesions in COVID-19 survivors: predictive clinical model

Carlos R R Carvalho (0000-0002-1618-8509)[1], Rodrigo C Chate[2], Marcio VY Sawamura[2], Michelle L Garcia[1], Celina A Lamas[1], Diego AC Cardenas[3], Daniel M Lima[3], Paula G Scudeller[1], João M Salge[1], Cesar H Nomura[2], Marco A Gutierrez[3], HCFMUSP Covid-19 Study Group*

1 Pulmonary Division, Heart Institute (InCor), Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo (HCFMUSP), Sao Paulo, SP, Brazil.

2 Radiology Institute (InRad), Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo (HCFMUSP), Sao Paulo, SP, Brazil.

3 Informatics Division, Heart Institute (InCor), Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo (HCFMUSP), Sao Paulo, SP, Brazil.

Correspondence to: Dr Carlos RR Carvalho (ORCID: 0000-0002-1618-8509), Pulmonary Division, Heart Institute (InCor), Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo (HCFMUSP), Av. Dr Eneas Carvalho de Aguiar, 44, Cerqueira Cesar, São Paulo, SP, 05403-900. Email: carlos.carvalho@hc.fm.usp.br. Fone: +55 11 26614505.

Word Count: 3331

Abstract

Objective This study aimed to propose a simple, accessible, and low-cost predictive clinical model to detect lung lesions due to COVID-19 infection.

Design, settings and participants: This prospective cohort study included COVID-19 survivors hospitalised between March 30, 2020 and August 31, 2020 followed-up after six months of discharge from a tertiary hospital in Sao Paulo, Brazil. There were 749 eligible RT-PCR-confirmed SARS-CoV-2 infected patients aged ≥18 years (median [IQR] age, 56 [44.4–65.1] years; 53% male). 257 patients had complete data and were included for the prediction analysis of pulmonary changes.

Outcome Measures: Anthropometric data and pulmonary function were assessed using the modified Medical Research Council (mMRC) dyspnoea scale, oximetry (SpO_2), spirometry (forced vital capacity [FVC]), and chest X-ray (CXR) during an in-person consultation. Patients with abnormalities in at least one of these parameters underwent chest computed tomography (CT). The median interval between hospital admission and consultation was 7.1 [6.7–8.5] months, and that between the first in-person consultation and chest CT was 45±33 days. mMRC scale, SpO_2, FVC, and CXR findings were used to build a machine learning model for lung lesion detection on CT.

Results There were 470 patients (63%) that had at least one sign of pulmonary involvement and were eligible for CT. 48% of them had significant pulmonary abnormalities, including ground-glass opacities, parenchymal bands, reticulation, traction bronchiectasis, and architectural distortion. The machine learning model accurately detected pulmonary lesions by the joint data of CXR, mMRC scale, SpO_2, and FVC (sensitivity, 0.85±0.08; specificity, 0.70±0.06; F1-score, 0.79±0.06; and AUC, 0.80±0.07).

Conclusion A predictive clinical model based on CXR, mMRC, oximetry, and spirometry data can accurately screen patients with lung lesions after SARS-CoV-2 infection. Given that these examinations are highly accessible and low cost, this protocol can be automated and implemented in different countries for early detection of COVID-19 sequelae.

Strengths and limitations of this study

- This study conducted a broad assessment, embracing an in-person clinical, functional, and radiological pulmonary examinations of a large cohort of COVID-19 patients.

- The sample size used for artificial intelligence evaluation was sufficient to provide a robust prediction equation.

- Although the study was conducted in a single centre, the cohort population was heterogeneous and hailed from all districts of the metropolitan region of Sao Paulo (with approximately 21 million inhabitants).

- Although there were some missing patient data and data lost to follow-up, in general they were from patients that had less severe disease.

INTRODUCTION

The coronavirus disease 2019 (COVID-19) caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) emerged in December 2019 and had since spread globally.[1] This multisystemic viral disease promotes endothelial and microvascular damage and immune system dysregulation, leading to hyperinflammatory and hypercoagulable states.[2 3] Several organs can be affected during the acute phase of COVID-19. In particular, pulmonary complications are considered life-threatening owing to the risk of progression to respiratory failure.[4 5]

COVID-19 symptoms can persist for more than 12 weeks after acute infection, characterizing long COVID.[1] The clinical complains of dyspnoea, fatigue, cough, chest pain, depression, cognitive disorders, headache, palpitations, myalgia, and arthralgia are the most reported in long COVID.[6-9] In addition to symptoms, some studies have shown that radiological abnormalities are also frequent in the follow-up of patients after the acute phase. In one of them, chest computed tomography (CT) was performed in 171 patients 4 months after hospital discharge and showed abnormalities in 75.5% of the patients who required invasive mechanical ventilation (IMV).[10] “Fibrotic-like changes” were observed in 19.3% of the total cohort and in 38.8% of patients with acute respiratory distress syndrome.[9] IMV can predict pulmonary sequelae, which reduce functional capacity and the health-related quality of life.[6 11 12] National Institute for Health and Care Excellence (NICE), has reported that some examinations can guide the diagnosis and management of post-COVID-19 syndrome,[1] including oximetry, spirometry, chest X-ray (CXR), ultrasonography, modified Medical Research Council (mMRC) dyspnoea scale, and chest CT. The latter examination is the gold standard for the diagnosis of chronic lung lesions due to COVID-19 and characterization of “fibrotic-like” lung lesions.[1 10]

The World Health Organization reported more than 265 million confirmed COVID-19 cases worldwide, with approximately 5 million deaths, and 260 million patients recovered as of December 2021.[13] The large number of recovered individuals experiencing long-term COVID-19 symptoms, such as fatigue, weakness, and dyspnoea, has drawn the attention of researches,[14 15] as they are expected to impose a significant health and economic burden.[14] In early 2021, the United Kingdom National Institute for Health Research invested £18.5 million to fund studies on long COVID.[16] The lack of knowledge and medical training for treating post-COVID symptoms also represents a significant public health challenge.[14] Thus, health care systems will have to reorganize themselves to address this issue, requiring the reallocation of resources and training of multidisciplinary teams and the development of new approaches.[14]

In this context, the wide availability of CXR and CT scanners has enabled the development of deep learning (DL) artificial intelligence-based algorithms for the automated diagnosis and prognosis of COVID-19.[17-19] For example, Castiglioni et al. [17] proposed a DL model for diagnosing COVID-19 with high sensitivity and specificity using radiography findings, whereas Wang et al. [18] developed a DL model (DenseNet) to classify CT images as positive or negative for COVID-19.

Although these works presented promising results, they were focused on images of patients in acute phase of COVID-19. However, as the pandemic is still ongoing with limited knowledge on long COVID-19 consequences,[20] a more comprehensive protocol for screening COVID-19 patients and assessing the risk of chronic pulmonary changes in recovered patients has not been validated to date. Thus, this study aimed to propose a predictive clinical model to detect the presence of radiologic chronic lung lesions due to SARS-CoV-2 infections based on data, including simple and accessible examinations, such as the mMRC dyspnoea scale, oximetry, spirometry, and CXR.

METHODS

Study design and eligibility

This prospective cohort study detected lung lesions in adult patients (≥ 18 years) with RT-PCR-confirmed SARS-CoV-2 infection admitted to the ward or

intensive care unit (ICU) of the Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo (HCFMUSP), Sao Paulo, Brazil, from March 30 to August 31st, 2020. It was considered only the first admission of each patient on the HCFMUSP. The protocols used in this study were described previously.[21] All research procedures were approved by the Research Ethics Committee of our institution (Process No. 31942020.0.000.0068).

The patients were invited to participate in the study six months after admission, and a face-to-face consultation was scheduled. At this point, all patients were already discharged. Clinical, radiological, and laboratory evaluations were performed after the patients gave written informed consent. Clinical data (comorbidities, cardiorespiratory symptoms, and smoking history), including the length of ICU stay and the need for IMV, were retrospectively collected from the electronic medical records of HCFMUSP. All data were stored in a structured form developed using REDCap software (https://www.redcapbrasil.com.br/).

General evaluation

All patients underwent a face-to-face consultation during the collection of anthropometric data and a pulmonary assessment, with an emphasis on respiratory symptoms. Dyspnoea was assessed using the mMRC scale.[21] Oxygen saturation (SpO_2) at rest and after physical exertion (1-min sit and stand test) was measured by pulse oximetry.[21 22] Spirometry was performed according to criteria established by ATS/ERS Task Force.[23] Actual spirometry results were compared with predicted values, according to Pereira et al. [24].

Then, patients underwent a posteroanterior and lateral CXR according to standard guidelines. The results of these examinations were evaluated blindly and independently by two chest radiologists (MVYS and RCC, have 7 and 16 years of experience in thoracic radiology, respectively) working on dedicated workstations. The radiographs were scored as 0 (results were normal or not related to COVID-19 [including cardiomegaly and pulmonary nodules, for instance]) or 1 (findings which could be related to COVID-19 [including bilateral linear and/or reticular opacities, especially peripheral opacities]). Disagreements were resolved by consensus. The agreement rate was 75%.

Previous classifications of radiographs by radiologists previously described were used to train and validate a DL algorithm to predict the probability that the CXR has findings related to COVID-19 sequelae. The DL algorithm is based on an EfficientNetB7 architecture[25] and a five-fold cross-validation strategy was adopted to train and validate the model, leading to an average area under the curve (AUC) of 0.89 (Supplemental Methods [Table 2]).

Chest CT

Patients who meet at least one the following criteria during the general evaluation were enrolled to undergo CT: (a) mMRC ≥ 2; (b) resting SpO_2 ≤ 90% and/or a decrease in SpO_2 of ≥4% during the 1-min sit and stand test; (c) opacities likely related to COVID-19 on CXR; and (d) FVC < lower limit of normal (LLN). The mean interval between CXR and chest CT was 45 ± 33 days.

The CT protocol used in this study was described previously.[21] CT findings consistent with COVID-19, including ground-glass and peripheral opacities, consolidations, parenchymal bands, reticulations, traction bronchiectasis, architectural distortions, honeycombing, bronchial wall thickening, mosaic attenuation, and pleural effusion, were categorized according to the criteria of the Fleischner Society.[26] The extent of lung involvement was quantified according to Francone et al. [27] by assigning the following scores to each pulmonary lobe: 0, none; 1, <5%; 2, 5-25%; 3, 26-50%; 4, 51-75%; and 5, >75%. The total score varied from 0 to 25 and was calculated by summing the scores of the five lobes. [25] Categorization of the CT features and score assignment were blindly and independently performed by the same two thoracic radiologists who evaluated the CXR (MVYS and RCC). Any disagreements were resolved by consensus.

A score ≥7 was used as the cut off value for significant CT changes after model calibration. The equations used to determine these scores are described in the Supplemental Methods.

Machine learning (ML) model

A Machine Learning (ML) model based on a Logistic Regression (LR) with L2 regularization to prevent overfitting[28] was adopted to detect the presence of

COVID-19-related chronic lung lesions. The L1 regularization was not included due to the variable selection by statistical significance that removed irrelevant and correlated attributes. In this ML model, the results of the mMRC scale, oximetry, spirometry, and DL-based classification of 257 CXR images were used as input data, and the presence of pulmonary lesions was used as output data (Figure 1). The performance of the model was evaluated by the metrics sensitivity, specificity, AUC, and F1-score after a five-fold cross validation. (Supplemental Methods)

Statistical analysis

Continuous variables are expressed as the mean and standard deviations or median and interquartile range. Normality of the variables was assessed by D'agostino-Pearson test. Normally and non-normally distributed continuous variables were compared using the Student's *t*-test and Mann-Whitney U test, respectively. Categorical variables are presented as counts and percentages and compared using the chi-square test. (Excel 2016; Python 3.8.11; extension packages: Pandas 1.0.1; Numpy 1.19.5; Scipy 1.5.4; Scikit-Learn 0.24.0).

The performance of the DL models was assessed by the area under the receiver operating characteristic (AUC) curve. The performance of the ML model was determined based on sensitivity, specificity, F1-score and AUC values (Supplemental Methods).

Patient and public involvement

Patients or the public were not involved in the design, conduct, reporting or dissemination plans of this research.

RESULTS

Of 3,753 COVID-19 enrolled patients, 1,957 were eligible for the study and 749 were included in the final analysis (445 [59%] and 304 [41%] patients were admitted to the ICU and ward, respectively). Additional information on the inclusion and exclusion criteria is shown in Figure 2.

Demographic characteristics of the cohort are shown in Supplemental Table S1. The median age was 56 years, with a predominance of overweight individuals, and 53% were male. Additionally, 59.4% of the patients were admitted to the ICU, and 68.5% of them were on IMV during the study period. The vital signs of most patients were within normal limits during the hospitalisation period (Supplemental Table S1).

The median interval between hospital admission and consultation was 7.1 (IQR [6.7–8.5]) months, and the lower and upper limits were 5.4 and 12.9 months, respectively. Of the 749 patients, 470 (63%) had at least one sign of pulmonary involvement (Table 1). Supplemental Figure S1 illustrates the simultaneous presence of two or more criteria for pulmonary involvement.

Table 1. Pulmonary function of patients with signs of pulmonary involvement (N=749).

Variables	Patients with signs of pulmonary involvement (N=749)
mMRC ≥ 2	229/742 (30.9)
Altered Oximetry*	71/675 (10.5)
CXR (score 1)	200/629 (31.8)
FVC < LLN	212/642 (33)

Values are presented as n/N (%). CXR, chest X-ray; FVC, forced vital capacity; mMRC, modified Medical Research Council dyspnoea scale; LLN, lower limit of normal. *Resting SpO_2 ≤90% or a decrease in SpO_2 of ≥4% during the 1-min sit and stand test.

The demographic and clinical characteristics of patients stratified by the presence of pulmonary involvement are described in Supplemental Table S2. Patients with pulmonary involvement were older and predominantly female, have more comorbidities, and a higher rate of ICU admission than those without (Supplemental Table S2). In patients with pulmonary involvement, 348 underwent CT (68%) (Figure 2). The demographic and clinical characteristics were similar between those that underwent or did not undergo the CT (Supplemental Table S3).

CT scores were obtained from 328 (94%) patients. Scores were not determined in 20 patients, who were excluded because of low CT scan quality or

had motion artefacts. Chest CT analysis showed that 47.6% of the patients had a score ≥7, and the most common features were ground-glass opacities, parenchymal bands, reticulation, traction bronchiectasis, and architectural distortions (Supplemental Table S4). In this group, 86.5% and 13.5% were admitted to the ICU and ward, respectively. Among the patients with normal CT findings (score = 0), 36.4% and 63.6% were admitted to the ICU and ward, respectively. The frequency of CT changes is shown in Supplemental Table S5. That frequency of "fibrotic-like" lesions, including traction bronchiectasis and architectural distortion, was significantly higher in the group admitted to the ICU in the acute phase of the disease. Long-term CT features in patients with moderate and critical COVID-19 are shown in Figure 3 and Supplemental Figure S2, respectively.

Of the 348 patients with CT data, 257 had data on mMRC, oximetry, spirometry, X-ray, and chest CT and were selected for the prediction analysis of pulmonary changes. Among the 91 patients excluded for the prediction analysis, 61 had incomplete data of all four tests (mMRC, oximetry, spirometry, CXR and CT) and 30 showed radiographic signs not related to COVID-19 (Supplemental Table S6).

Three data groups were considered for the prediction analysis of pulmonary changes: (1) clinical data (oximetry [SpO_2], mMRC dyspnoea scores, and spirometry [FVC]), (2) CXR, and (3) all variables (oximetry [SpO_2], mMRC dyspnoea scores, spirometry [FVC], and CXR). The performance of the predictive model was higher using the combination of all variables (clinical variables and CXR), and the following metrics expressed in terms of mean ± standard deviation and 95% Confidence Interval (CI) were considered: sensitivity, 0.85±0.08 (95% CI [0.77, 0.94]); specificity, 0.70±0.14 (95% CI [0.55, 0.85]); F1-score, 0.79±0.06 (95% CI [0.73, 0.85]); and AUC, 0.80±0.07(95% CI [0.72, 0.87]) (Table 2).

Table 2. Performance of the predictive model using three combinations of variables (N=257).

Groups of variables	Sensitivity	Specificity	F1-score	AUC
1 SpO_2, mMRC score, and FVC	0.87±0.16	0.42±0.33	0.71±0.03	0.68±0.10
2 CXR	0.88±0.05	0.52±0.14	0.75±0.04	0.78±0.05
3 SpO_2, mMRC score, FVC, and CXR	0.85±0.08	0.70±0.14	0.79±0.06	0.80±0.07

Values are presented as means ± standard deviations after five-fold cross validation for each test fold. CXR, chest X-Ray; FVC, forced vital capacity; mMRC, Modified Medical Research Council dyspnoea scale.

The machine learning predictive model is represented by the following function:

$$p_{CT} = \beta_1 FVC^* + \beta_2\, mMRC^* + \beta_3 S_p O_2 + \beta_4 p_{CXR0} + \beta_5 p_{CXR1} + \beta_6 p_{CXR2} + \beta_7 p_{CXR3} + \beta_8 p_{CXR4}$$

$$\beta_1 = -0.3705 \;\; \beta_2 = -2.2807 \;\; \beta_3 = -0.745 \;\; \beta_4 = 1.1257$$

$$\beta_5 = 1.4960 \;\; \beta_6 = 1.0761 \;\; \beta_7 = 0.7328 \;\; \beta_8 = -0.7613$$

where p_{CT} is the probability of the presence of abnormalities on CT images, $FVC^* = \frac{FVC_{Resting}}{2FVC_{min}}$, $mMRC^* = \frac{mMRC}{4}$, and p_{CXR0} to p_{CXR4} are the probabilities that the CXR image has findings related to sequelae from COVID-19, obtained in each fold (0 to 4) during a 5-folds cross validation. (Supplemental Methods)

Therefore, based in these observations, we propose in a flowchart a suggestion for lung lesion case-finding in COVID-19 survivors (Figure 4).

DISCUSSION

Few studies have assessed the pulmonary abnormalities in COVID-19 survivors after six months of hospital discharge. However, some of these patients

have developed long-term pulmonary complications after the acute phase of the disease.[6 29-33] This study evaluated 749 COVID-19 patients who received supplemental oxygen or ventilatory support in the ward or ICU and survived. They underwent an in-person comprehensive clinical, functional, and radiological assessments, which were more extensive than those performed in previous studies,[6 30 31 33-35] conferring reliability to our results.

In the first months after recovery, the most common CT findings in COVID-19 hospitalised patients included ground-glass opacities, parenchymal bands, reticulation, mosaic attenuation pattern, and "fibrotic-like" abnormalities, including traction bronchiectasis and architectural distortions.[36 37] These findings were detected in 76.5% of our cohort, and severe and extensive changes were noted in approximately 50% of the cases. The CT abnormalities were more prevalent in older critical patients and individuals with more comorbidities, which is consistent with previous studies.[32 38] These results indicate the high prevalence of chronic lung lesions and sequelae in post-COVID patients worldwide.

Therefore, the need to identify severe pulmonary complications due to COVID-19, including fibrosis,[1] and the large number of COVID-19 survivors, prompted us to develop a predictive clinical model to screen patients admitted to a tertiary hospital, which could be able to reduce costs and radiation exposure. During the first six months of the pandemic in Sao Paulo, Brazil, all hospital beds at HCFMUSP (300 in the ICU and 400 in the ward) were made available to COVID-19 patients.[12] Patients were treated free of charge in our hospital owing to a universal health system, and there is a constant search for better and cost-effective protocols to improve workflow.[12]

Dyspnoea scales, CXR, oximetry, and spirometry are commonly used to evaluate COVID-19 symptoms.[2] A Norwegian study evaluated a cohort of 100 patients three months after admission to a hospital and reported that 19% had dyspnoea (mMRC score>1) and 10% presented altered FVC and normal oxygen saturation levels, suggesting the lower sensitivity of pulse oximetry.[39] In 113 patients evaluated 4 months after COVID-19 diagnosis in Switzerland, FVC and oxygen saturation levels were lower in patients who had a severe disease than in those with a moderate disease, although the mean values remained within the

limits of normality.[35] In addition, a previous study has suggested that cough, lymphocytosis and the lung volume could indicate lung lesions in COVID-19-recovered patients.[34]

Ground-glass and reticular opacities can be detected by CXR, although this method is less sensitive than CT.[40] On the other hand, CXR is readily available in the primary care setting and has a lower cost and radiation exposure than CT.[40 41] Radiographs were separately scored by an automated DL-based image analysis tool and chest radiology specialists, and there was a high level of consensus between these scores (AUC = 0.89). In the Brazilian public health system, the cost of a CT scan is approximately 15 times higher than that of a CXR.[41] According to the American College of Radiology and the Radiological Society of North American, the radiation doses of a standard chest CT and CXR are 6.1 mSv and 0.1 mSv, respectively; this underscores the advantage of CXR in reducing the exposure of COVID-19 patients to radiation, especially those who have already performed serial imaging exams in the acute phase of the disease.[42]

Nevertheless, none of these examinations alone accurately predicted pulmonary complications. The performance of our model corroborates this finding since the information provided by each clinical examination alone did not accurately diagnose the pulmonary changes detected on CT. In contrast, clinical and radiographic data were complementary and increased the performance of the ML model. Cross-validation also increased the robustness of the results. These results indicate that four examinations (oximetry, mMRC dyspnoea scale, spirometry, and CXR) should be jointly conducted to screen patients at risk of developing chronic lung lesions due to COVID-19 and achieve a diagnostic performance similar to that of CT (sensitivity, 0.85±0.08; specificity, 0.70±0.14; F1-score, 0.79±0.06; and AUC, 0.80±0.07). Analysis of these metrics indicates that this predictive clinical method can better identify the true positives than true negatives. In addition, the F1-score takes into account both false-positive and false-negative results and measures the accuracy of the method in the dataset.

The WHO has highlighted the importance of establishing screening protocols with a favourable cost-effectiveness ratio for patients affected by different pathologies.[43] The identification of COVID-19 lung lesions will allow the

accurate referral of patients to specialists for further investigation and treatment. As the COVID-19 sequelae can progress to increasing intensity of symptoms and risk of disability, this approach can improve the quality and length of life of patients, since medical interventions can be performed as early as possible.

We already have an initiative to implement this protocol in Brazil. The project will start in the state of Sao Paulo, in partnership with the State of Sao Paulo Health Department, where the HCFMUSP is located. We will start to apply this screening protocol in the central area of the city of Sao Paulo, with approximately 430.000 inhabitants, according to the flowchart suggested for lung lesion case-finding in COVID-19 survivors (Figure 4). Firstly, exams will be performed in the following order, starting from the simplest and most accessible ones: oximetry/mMRC, spirometry and CXR. At the moment the patient shows alterations in any of these four exams, the patient will be enrolled directly for further investigation in a specialised care centre to perform CT and/or other specific exams. We expect that over time, this can lead to a significant reduction in morbidity and mortality due to COVID-19 lung sequelae, relieving the burden on the health care system, reducing expenses of imaging exams and accelerating the medical interventions.

This study has some limitations. First, there was variability in the interval between the execution of CXR and CT. Notwithstanding this variation, which might contribute to lung recovery, our protocol screened a large number of patients with pulmonary lesions, demonstrating the persistence of these manifestations secondary to COVID-19 and reducing sampling bias. Second, the single-centre nature of the study limits the generalizability of our results. However, a previous study showed that the population of patients admitted to HCFMUSP—a tertiary reference hospital for the treatment of COVID-19 in Brazil—was heterogeneous and hailed from all districts of the metropolitan region of Sao Paulo (with approximately 21 million inhabitants).[12] Third, we were unable to contact some patients because of inconsistencies in telephone numbers and addresses. Thus, these subjects were not included in the protocol, although public death registry data showed that they were alive. Fourth, this screening protocol was developed based on respiratory complaints, which are considered

risk factors for the development of chronic lung complications. However, other COVID-19 symptoms were not analysed in this study.

The breadth of our results allowed us to propose a simple, accessible, and low-cost clinical predictive model to screen patients at risk of developing chronic lung lesions due to COVID-19. The low cost and easy accessibility to these examinations facilitate the implementation of the proposed protocol in developing countries. In addition, it may contribute to early and effective determination of the treatment course, thus reducing radiation exposure and the conduct of costly imaging examinations. The use of artificial intelligence facilitated the large-scale assessment of radiographs and their association with clinical variables, demonstrating that artificial intelligence models can be used to automate diagnosis, especially in severe patients.

Collaborators: *Members of the HCFMUSP Covid-19 Study Group: Adriana L Araújo, Aluisio C Segurado, Amanda C Montal, Anna Miethke-Morais, Anna S Levin, Beatriz Perondi, Bruno F Guedes, Carolina Carmo, Carolina S Lázari, Cassiano C Antonio, Clarice Tanaka, Claudia C Leite, Cristiano Gomes, Edivaldo M Utiyama, Emmanuel A Burdmann, Eloisa Bonfá, Esper G Kallas, Ester Sabino, Euripedes C Miguel, Fabio R Pinna, Fabiane Y O Kawano, Geraldo F Busatto, Giovanni G Cerri, Guilherme Fonseca, Heraldo P Souza, Izabel Marcilio, Izabel C Rios, Jorge Hallak, José Eduardo Krieger, Juliana C Ferreira, Julio F M Marchini, Larissa S Oliveira, Leila Harima, Linamara R Batisttella, Luis Yu, Luiz Henrique M Castro, Marcelo C Rocha , Marcello M C Magri, Marcio Mancini, Maria Amélia de Jesus, Maria Cassia J M Corrêa, Maria Cristina P B Francisco, Maria Elizabeth Rossi, Marjorie F Silva, Marta Imamura, Maura S Oliveira, Nelson Gouveia, Orestes V Forlenza, Paulo A Lotufo, Ricardo F Bento, Ricardo Nitrini, Rodolfo F Damiano, Roger Chammas, Rossana P Francisco, Solange R G Fusco, Tarcisio E P Barros-Filho, Thais Mauad, Thaís Guimarães, Thiago Avelino-Silva and Wilson J Filho.

Funding: This study was funded by the Sao Paulo Research Foundation (grant number - 2020/07200-9).

Contributors: CRRC: conceptualisation, data curation, formal analysis, funding acquisition, investigation, methodology, project administration, resources, supervision, validation, visualisation, writing – original draft, and writing – review & editing. RCC: data curation, formal analysis, investigation, methodology, validation, visualisation, writing – original draft, and writing – review & editing. MVYS: data curation, formal analysis, investigation, methodology, validation, visualisation, writing – original draft, and writing – review & editing. MLG: conceptualisation, data curation, formal analysis, investigation, methodology, project administration, supervision, validation, visualisation, writing – original draft, and writing – review & editing. CAL: data curation, formal analysis, investigation, writing – original draft, and writing – review & editing. DACC: data curation, formal analysis, methodology, software, writing – original draft, and writing – review & editing. DML: data curation, formal analysis, methodology, software, writing – original draft, and writing – review & editing. PGS: conceptualisation, project administration, supervision, validation, visualisation and writing – review & editing. JMS: methodology, validation, visualisation and writing – review & editing. CHN: methodology, validation, visualisation and writing – review & editing. MAG: data curation, formal analysis, funding acquisition, methodology, software, supervision, validation, visualisation, writing – original draft, and writing – review & editing. HCFMUSP Covid-19 Study Group: contributed to the implementation of the study and data collection. All authors critically reviewed and approved the final version.

Competing interests' statement: None declared.

Data Sharing: The study protocol was previously described by Busatto et al. [21] and was registered at the "Brazilian Registry of Clinical Trials" (https://ensaiosclinicos.gov.br/). The raw data are not publicly available because follow-up studies will be carried out. However, data are available from the corresponding author upon request and authorization from the institution. Data on demographics, hospitalisation, and outcomes are available in the COVID-19 Data Sharing/BR repository and are freely available for download[44].

Ethical approval: The study was approved by the Research Ethics Committee of the HCFMUSP (approval number 31942020.0.000.0068).

REFERENCES

1. Sisó-Almirall A, Brito-Zerón P, Conangla Ferrín L, et al. Long Covid-19: Proposed Primary Care Clinical Guidelines for Diagnosis and Disease Management. *Int J Environ Res Public Health* 2021;18(8) doi: 10.3390/ijerph18084350 [published Online First: 2021/04/20]
2. Nalbandian A, Sehgal K, Gupta A, et al. Post-acute COVID-19 syndrome. *Nat Med* 2021;27(4):601-15. doi: 10.1038/s41591-021-01283-z [published Online First: 2021/03/22]
3. Mauad T, Duarte-Neto AN, da Silva LFF, et al. Tracking the time course of pathological patterns of lung injury in severe COVID-19. *Respir Res* 2021;22(1):32. doi: 10.1186/s12931-021-01628-9 [published Online First: 20210129]
4. Tanni SE, Fabro AT, de Albuquerque A, et al. Pulmonary fibrosis secondary to COVID-19: a narrative review. *Expert Rev Respir Med* 2021;15(6):791-803. doi: 10.1080/17476348.2021.1916472 [published Online First: 2021/04/27]
5. Macedo BR, Garcia MVF, Garcia ML, et al. Implementation of Tele-ICU during the COVID-19 pandemic. *J Bras Pneumol* 2021;47(2):e20200545. doi: 10.36416/1806-3756/e20200545 [published Online First: 20210430]
6. Huang C, Huang L, Wang Y, et al. 6-month consequences of COVID-19 in patients discharged from hospital: a cohort study. *Lancet* 2021;397(10270):220-32. doi: 10.1016/S0140-6736(20)32656-8 [published Online First: 2021/01/08]
7. Lopez-Leon S, Wegman-Ostrosky T, Perelman C, et al. More Than 50 Long-Term Effects of COVID-19: A Systematic Review and Meta-Analysis. *Res Sq* 2021 doi: 10.21203/rs.3.rs-266574/v1 [published Online First: 2021/03/01]
8. Fernandes PMP, Mariani AW. Life post-COVID-19: symptoms and chronic complications. *Sao Paulo Med J* 2021;139(1):1-2. doi: 10.1590/1516-3180.2021.139104022021
9. Morin L, Savale L, Pham T, et al. Four-Month Clinical Status of a Cohort of Patients After Hospitalization for COVID-19. *JAMA* 2021;325(15):1525-34. doi: 10.1001/jama.2021.3331
10. Raghu G, Collard HR, Egan JJ, et al. An official ATS/ERS/JRS/ALAT statement: idiopathic pulmonary fibrosis: evidence-based guidelines for diagnosis and management. *Am J Respir Crit Care Med* 2011;183(6):788-824. doi: 10.1164/rccm.2009-040GL
11. Carfì A, Bernabei R, Landi F, et al. Persistent Symptoms in Patients After Acute COVID-19. *JAMA* 2020;324(6):603-05. doi: 10.1001/jama.2020.12603
12. Ferreira JC, Ho YL, Besen BAMP, et al. Protective ventilation and outcomes of critically ill patients with COVID-19: a cohort study. *Ann Intensive Care* 2021;11(1):92. doi: 10.1186/s13613-021-00882-w [published Online First: 2021/06/07]
13. WHO. WHO Coronavirus Disease (COVID-19) Dashboard 2021 [cited 2021 October 13]. Available from: https://covid19.who.int/ accessed October 13 2021
14. Wade DT. Rehabilitation after COVID-19: an evidence-based approach. *Clin Med (Lond)* 2020;20(4):359-65. doi: 10.7861/clinmed.2020-0353 [published Online First: 2020/06/09]
15. Godoy CG, Silva ECGE, Oliveira DB, et al. Protocol for Functional Assessment of Adults and Older Adults after Hospitalization for COVID-19. *Clinics (Sao Paulo)* 2021;76:e3030. doi: 10.6061/clinics/2021/e3030 [published Online First: 20210614]

16. Subbaraman N. US health agency will invest $1 billion to investigate 'long COVID'. *Nature* 2021;591(7850):356. doi: 10.1038/d41586-021-00586-y
17. Castiglioni I, Ippolito D, Interlenghi M, et al. Machine learning applied on chest x-ray can aid in the diagnosis of COVID-19: a first experience from Lombardy, Italy. *Eur Radiol Exp* 2021;5(1):7. doi: 10.1186/s41747-020-00203-z [published Online First: 20210202]
18. Wang S, Zha Y, Li W, et al. A fully automatic deep learning system for COVID-19 diagnostic and prognostic analysis. *Eur Respir J* 2020;56(2) doi: 10.1183/13993003.00775-2020 [published Online First: 20200806]
19. Ferreira Junior JR, Cardona Cardenas DA, Moreno RA, et al. Novel Chest Radiographic Biomarkers for COVID-19 Using Radiomic Features Associated with Diagnostics and Outcomes. *J Digit Imaging* 2021;34(2):297-307. doi: 10.1007/s10278-021-00421-w [published Online First: 20210218]
20. Greenhalgh T, Knight M, A'Court C, et al. Management of post-acute covid-19 in primary care. *BMJ* 2020;370:m3026. doi: 10.1136/bmj.m3026 [published Online First: 2020/08/11]
21. Busatto GF, de Araújo AL, Duarte AJDS, et al. Post-acute sequelae of SARS-CoV-2 infection (PASC): a protocol for a multidisciplinary prospective observational evaluation of a cohort of patients surviving hospitalisation in Sao Paulo, Brazil. *BMJ Open* 2021;11(6):e051706. doi: 10.1136/bmjopen-2021-051706 [published Online First: 2021/06/30]
22. van den Borst B, Peters JB, Brink M, et al. Comprehensive health assessment three months after recovery from acute COVID-19. *Clin Infect Dis* 2020 doi: 10.1093/cid/ciaa1750 [published Online First: 2020/11/21]
23. Miller MR, Hankinson J, Brusasco V, et al. Standardisation of spirometry. *Eur Respir J* 2005;26(2):319-38. doi: 10.1183/09031936.05.00034805
24. Pereira CA, Sato T, Rodrigues SC. New reference values for forced spirometry in white adults in Brazil. *J Bras Pneumol* 2007;33(4):397-406. doi: 10.1590/s1806-37132007000400008
25. Tan M, Le Q. Efficientnet: Rethinking model scaling for convolutional neural networks. *International Conference on Machine Learning* 2019:6105-14.
26. Hansell DM, Bankier AA, MacMahon H, et al. Fleischner Society: glossary of terms for thoracic imaging. *Radiology* 2008;246(3):697-722. doi: 10.1148/radiol.2462070712 [published Online First: 20080114]
27. Francone M, Iafrate F, Masci GM, et al. Chest CT score in COVID-19 patients: correlation with disease severity and short-term prognosis. *Eur Radiol* 2020;30(12):6808-17. doi: 10.1007/s00330-020-07033-y [published Online First: 2020/07/04]
28. Vittinghoff E, Glidden DV, Shiboski SC, et al. Regression Methods in Biostatistics: Linear, Logistic, Survival, and Repeated Measures Models. 2nd ed. New York: Springer-Verlag, 2012:1272.
29. Wu Q, Zhong L, Li H, et al. A Follow-Up Study of Lung Function and Chest Computed Tomography at 6 Months after Discharge in Patients with Coronavirus Disease 2019. *Can Respir J* 2021;2021:6692409. doi: 10.1155/2021/6692409 [published Online First: 20210213]
30. Huang L, Yao Q, Gu X, et al. 1-year outcomes in hospital survivors with COVID-19: a longitudinal cohort study. *Lancet* 2021;398(10302):747-58. doi: 10.1016/S0140-6736(21)01755-4
31. Han X, Fan Y, Alwalid O, et al. Fibrotic Interstitial Lung Abnormalities at 1-year Follow-up CT after Severe COVID-19. *Radiology* 2021:210972. doi: 10.1148/radiol.2021210972 [published Online First: 20210727]
32. Han X, Fan Y, Alwalid O, et al. Six-month Follow-up Chest CT Findings after Severe COVID-19 Pneumonia. *Radiology* 2021;299(1):E177-E86. doi: 10.1148/radiol.2021203153 [published Online First: 20210126]

33. Stylemans D, Smet J, Hanon S, et al. Evolution of lung function and chest CT 6 months after COVID-19 pneumonia: Real-life data from a Belgian University Hospital. *Respir Med* 2021;182:106421. doi: 10.1016/j.rmed.2021.106421 [published Online First: 20210418]
34. Caruso D, Guido G, Zerunian M, et al. Post-Acute Sequelae of COVID-19 Pneumonia: Six-month Chest CT Follow-up. *Radiology* 2021;301(2):E396-E405. doi: 10.1148/radiol.2021210834 [published Online First: 20210727]
35. Guler SA, Ebner L, Aubry-Beigelman C, et al. Pulmonary function and radiological features 4 months after COVID-19: first results from the national prospective observational Swiss COVID-19 lung study. *Eur Respir J* 2021;57(4) doi: 10.1183/13993003.03690-2020 [published Online First: 2021/04/29]
36. Liu C, Ye L, Xia R, et al. Chest Computed Tomography and Clinical Follow-Up of Discharged Patients with COVID-19 in Wenzhou City, Zhejiang, China. *Ann Am Thorac Soc* 2020;17(10):1231-37. doi: 10.1513/AnnalsATS.202004-324OC
37. Tabatabaei SMH, Rajebi H, Moghaddas F, et al. Chest CT in COVID-19 pneumonia: what are the findings in mid-term follow-up? *Emerg Radiol* 2020;27(6):711-19. doi: 10.1007/s10140-020-01869-z [published Online First: 20201109]
38. Solomon JJ, Heyman B, Ko JP, et al. CT of Postacute Lung Complications of COVID-19. *Radiology* 2021:211396. doi: 10.1148/radiol.2021211396 [published Online First: 20210810]
39. Sonnweber T, Sahanic S, Pizzini A, et al. Cardiopulmonary recovery after COVID-19: an observational prospective multicentre trial. *Eur Respir J* 2021;57(4) doi: 10.1183/13993003.03481-2020 [published Online First: 2021/04/29]
40. Jacobi A, Chung M, Bernheim A, et al. Portable chest X-ray in coronavirus disease-19 (COVID-19): A pictorial review. *Clin Imaging* 2020;64:35-42. doi: 10.1016/j.clinimag.2020.04.001 [published Online First: 20200408]
41. DataSUS. MdS. SIGTAP - Sistema de Gerenciamento da Tabela de Procedimentos, Medicamentos e OPM do SUS. 2021 [cited 2021 January 03]. Available from: http://sigtap.datasus.gov.br/tabela-unificada/app/sec/inicio.jsp accessed January 03 2021.
42. Mettler FAM, Mahadevappa Bhargavan, MythreyiChambers, Charles E.Elee, Jennifer G. Frush, Donald P. Milano, Michael T. Miller, Donald L. Royal, Henry D. Spelic, David C.Ansari, Armin J. Bolch, Wesley E.Guebert, Gary M. Sherrier, Robert H. Smith, James M. Vetter, Richard J. Report Na. 184 - Medical Radiation Exposure of Patients in the United States. United States: National Council on Radiation Protection and Measurements, 2019.
43. Wilson J, Jungner G. Principles and Practice of Screening for Disease.: World Health Organization Public Health Papers, 1968.
44. FAPESP. COVID-19 DataSharing/BR 2021 [cited 2021 September 08]. Available from: https://repositoriodatasharingfapesp.uspdigital.usp.br/.

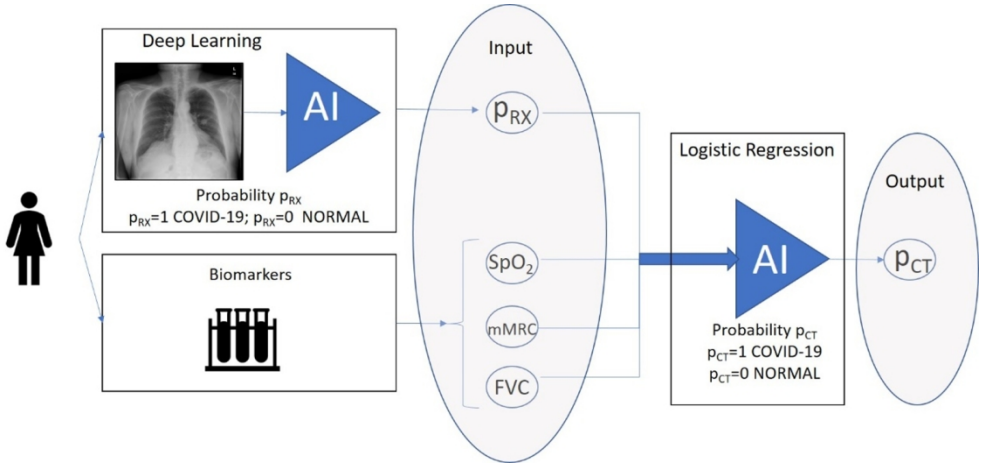


Figure 1. Logistic regression-based machine learning model. The modified Medical Research Council (mMRC) dyspnoea scale, oximetry (SpO2), spirometry (forced vital capacity [FVC]), and the five radiographic scores obtained during DL-based classification of CXR (pRX) were used as input data, and the presence of CT lung lesions due to COVID-19 was used as output data. AI, artificial intelligence; CT, computed tomography.

68x31mm (600 x 600 DPI)

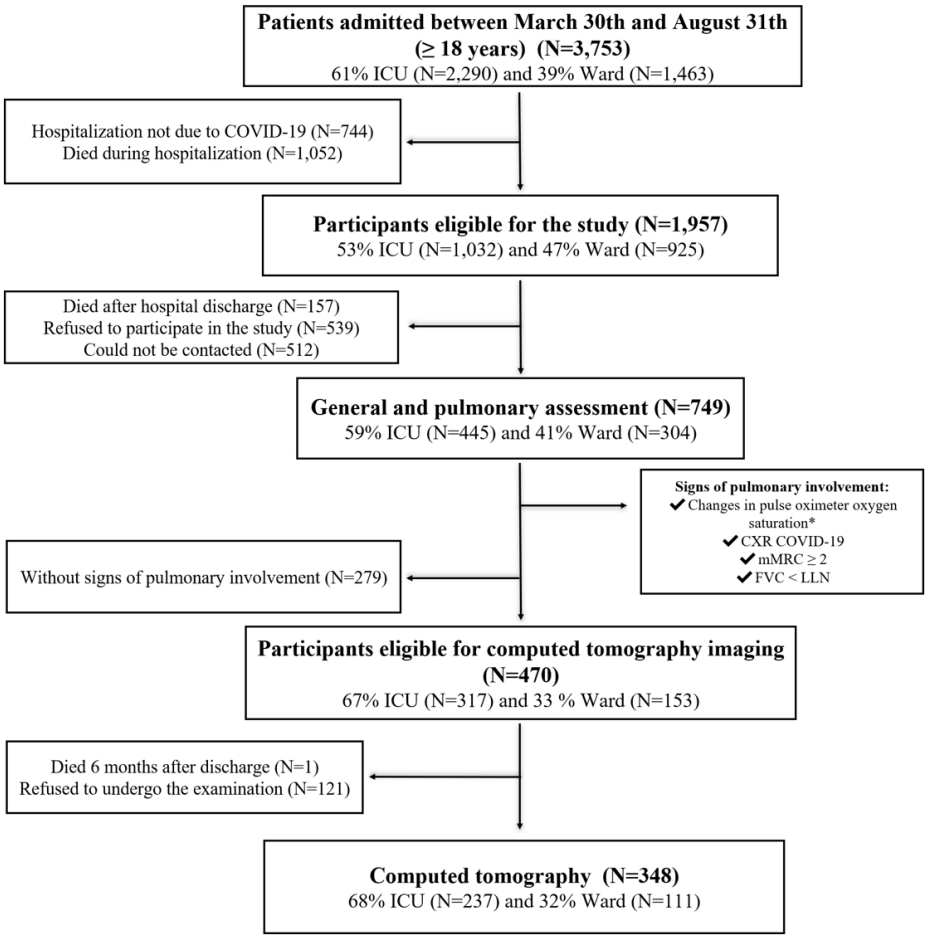


Figure 2. Flowchart of patient selection. FVC, forced vital capacity; LLN, lower limit of normal; mMRC, modified Medical Research Council dyspnoea scale. *Rest SpO2 < 90% or a decrease in SpO2 of at least 4% after the 1-min sit and stand test.

51x47mm (600 x 600 DPI)

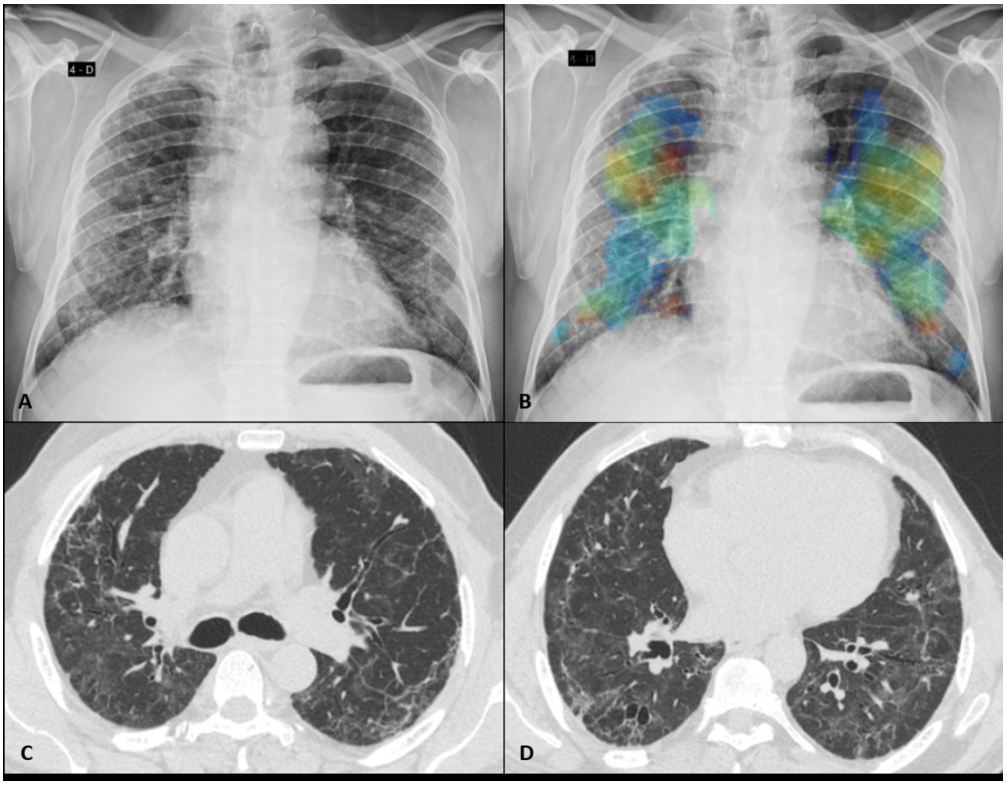


Figure 3. Fibrotic-like changes after critical COVID-19 in a patient in his early 70s. (A) PA chest radiograph obtained 7 months after infection shows reticular opacities with a slight peripheral predominance diffusely distributed in both lungs. (B) Image from the same radiograph analysed by the AI algorithm with a heat map highlighting the areas of pulmonary involvement. (C, D) Chest CT obtained 8 months after infection shows moderate ground glass opacities, linear multifocal and reticular abnormalities, discrete traction bronchiectasis and slight parenchymal architectural distortion. The patient had dyspnoea (mMRC=1) and altered FVC (2.34 L / 60% pred), besides the normal oximetry (97%).

154x119mm (600 x 600 DPI)

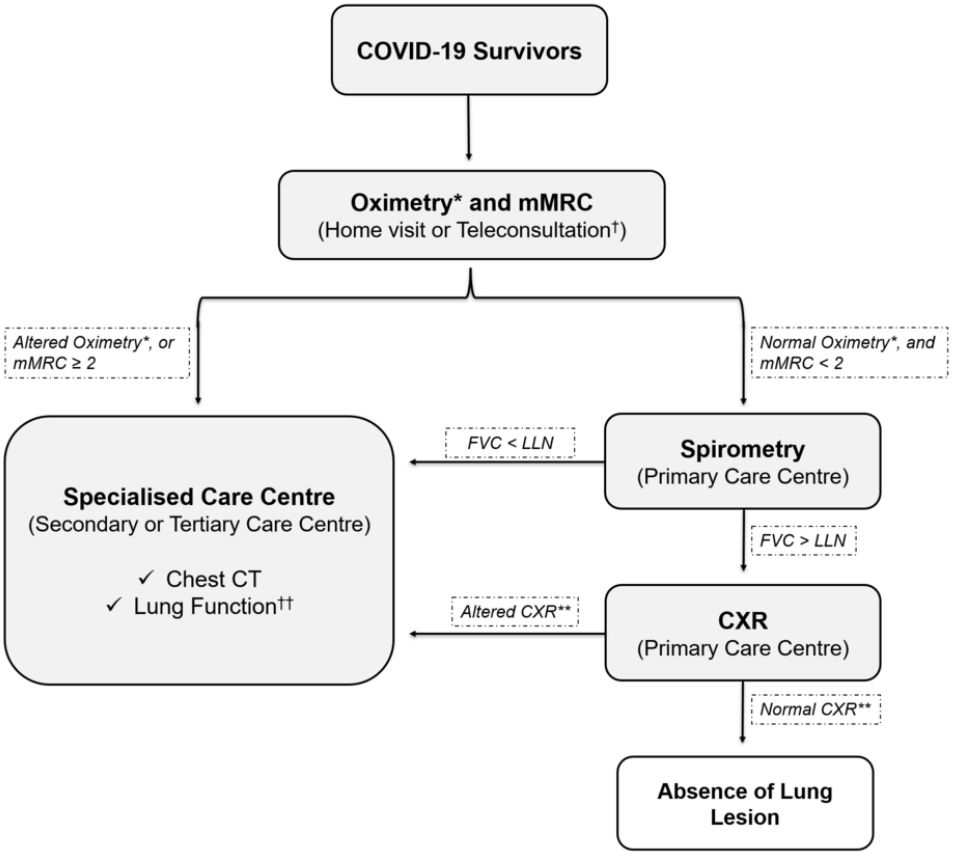


Figure 4. Flowchart for lung lesion case-finding in COVID-19 survivors. *Altered oximetry: Resting SpO2 ≤90% or a decrease in SpO2 of ≥4% during the 1-min sit and stand test. **Altered CXR: COVID-19 findings, including bilateral linear and/or reticular opacities, especially peripheral opacities. † The in-person consultation also should start with oximetry and mMRC examinations. †† The suggestion is to perform plethysmography with diffusion capacity measure. CXR, chest X-Ray; FVC, forced vital capacity; LLN, lower limit of normal; mMRC, modified Medical Research Council dyspnea scale.

81x71mm (300 x 300 DPI)

Supplemental Material

Chronic lung lesions in COVID-19 survivors: predictive clinical model

Carlos R R Carvalho, Rodrigo C Chate, Marcio VY Sawamura, Michelle L Garcia, Celina A Lamas, Diego AC Cardenas, Daniel M Lima, Paula G Scudeller, João M Salge, Cesar H Nomura, Marco A Gutierrez, HCFMUSP Covid-19 Study Group.

Contents

Supplemental Methods

Datasets

The SIIM-RSNA dataset contains 6,334 posterior-anterior radiographic images from 6,054 patients obtained from the public dataset Machine Learning Challenge on COVID-19 Pneumonia Detection and Localization.[1] Specialists classified images as “negative for pneumonia” or “COVID-19 pneumonia”. A total of 6,030 images were selected and randomly distributed in training and validation sets (1,276 negative and 3,711 positive findings) and a test set (400 negative and 643 positive findings).

The Institute of Radiology (InRad) dataset contains chest X-Ray (CXR) and chest computed tomographic (CT) images of 257 patients. The CXR images were classified as normal (n=145) or with findings related to COVID-19 (n=112) and randomly distributed in training and validation sets (214 patients) and a test set (n=43). Images were obtained from the InRad of the Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo (HCFMUSP).

Because of differences in dataset sizes, a data augmentation technique was adopted using random transformations, including rotation (0–15 degrees), horizontal mirroring, and random changes in intensity and contrast (0–5%).

Classification of chest radiography images

A deep-learning (DL) approach using a convolutional neural network (CNN) based on an EfficientNetB7 architecture was used.[2] The network classification layer was replaced by a global average pooling operation, followed by batch normalization and the adoption of a dense layer with one neuron and sigmoid activation function. Each training iteration was run for 40 epochs with an Adam optimizer at a learning rate of 0.0001. All images were resized to 600 x 600 pixels.

The CNN was trained using the SIIM-RSNA dataset to detect radiographic patterns of COVID-19 pneumonia. Training was initiated in EfficientNetB7 using weights after pre-training with the ImageNet dataset.[3]

A five-fold cross-validation strategy was adopted for the training and validation sets. The training weights obtained for each fold were used with the

test set of the SIIM-SNA to evaluate classification accuracy (Table 1). The fold with the best area under the receiver operating characteristic curve (AUC), in this case, fold 1 with AUC of 0.89, defines the final weights of the CNN.

Table 1. Classification of the test set of the SIIM-RSNA dataset as negative (normal) or positive (patterns of COVID-19 pneumonia).

Dataset	5-fold	Acc	Prec	Sensitivity	Specificity	F1-score	AUC
SIIM-RSNA	0	0.80	0.85	0.82	0.76	0.83	0.88
	1	0.80	0.85	0.82	0.77	0.84	**0.89**
	2	0.78	0.77	0.92	0.56	0.84	0.87
	3	0.76	0.74	0.93	0.48	0.83	0.86
	4	0.76	0.74	0.93	0.48	0.83	0.86

Area under the receiver operating characteristic curve (AUC); Accuracy (Acc); Precision (Prec).

For the InRad dataset, the CNN was initialized with the final weights defined in the training set of SIIM-RSNA. After initialization, the CNN was retrained to classify images as normal or with findings related to COVID-19.

The InRad dataset was divided into six-folds during the retraining, five folds for training and validation, and one-fold for test. To avoid bias, the test fold was selected to run all six folds available and, for each test fold selected, a five-fold cross-validation strategy was applied in the remaining training and validation folds (Table 2).

Table 2. Classification using six test folds of the InRad database.

Dataset	Test fold	Acc	Prec	Sensitivity	Specificity	F1-score	AUC
InRad	0	0.79±0.01	0.74±0.04	0.82±0.07	0.77±0.06	0.78±0.02	0.86±0.02
	1	0.69±0.02	0.62±0.03	0.84±0.06	0.57±0.07	0.71±0.02	0.75±0.01
	2	0.67±0.05	0.60±0.06	0.81±0.08	0.57±0.13	0.68±0.02	0.76±0.02
	3	0.77±0.04	0.71±0.07	0.80±0.04	0.74±0.10	0.75±0.03	0.80±0.02
	4	0.82±0.05	0.77±0.11	0.89±0.10	0.78±0.14	0.81±0.03	0.89±0.04
	5	0.71±0.04	0.62±0.04	0.90±0.02	0.58±0.08	0.73±0.03	0.80±0.02

Data represent the mean and standard deviation after five-fold cross validation. Area under the receiver operating characteristic curve (AUC); Accuracy (Acc); Precision (Prec).

Detection of chronic lung lesions on computed tomography images

Three machine learning models were developed based on the clinical data, including the modified Medical Research Council dyspnea scale (mMRC), oximetry (SpO_2) and spirometry (forced vital capacity, FVC), and five radiographic

probabilities (p_{RX0} to p_{RX4}) with findings related to COVID-19 (p_{RXn}=1) and normal (p_{RXn}=0), which were obtained from the previous step (Table 2). As output, the models predict the value of a binary variable (p_{CT}) related to the presence of chronic lung lesions on CT images, with p_{CT}=1 for a CT score ≥ 7 (n=129) and p_{CT}=0 for a CT score < 7 (n=128) (Figure 1).

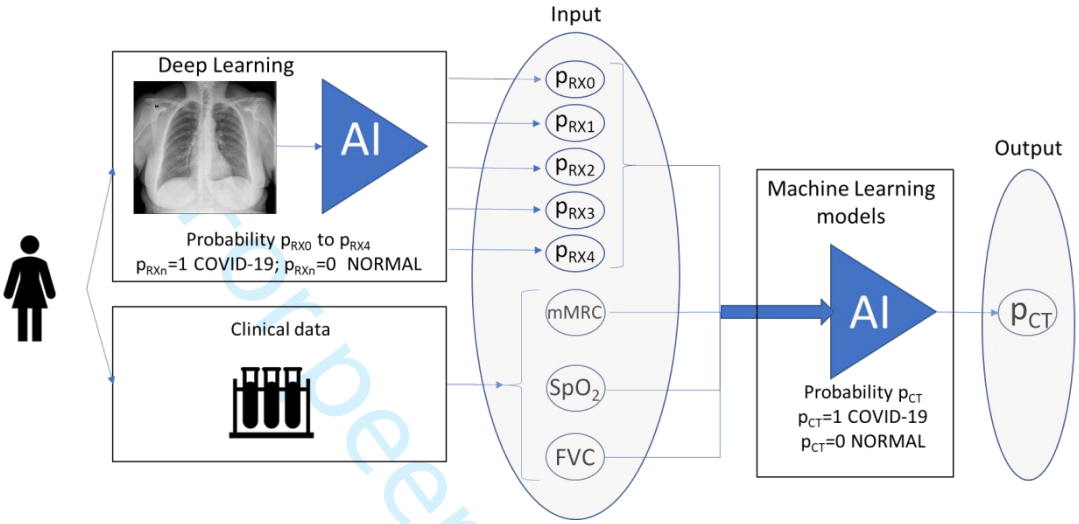


Figure 1. Machine learning-based model. Data on the modified Medical Research Council (mMRC) dyspnea scale, oximetry (SpO2), and spirometry (forced vital capacity [FVC]), and radiographic probabilities (p_{RX0} to p_{RX4}) with findings related to COVID-19 (pRXn=1) and normal (pRXn=0) were used as input variables, and the presence of lung lesions due to COVID-19 (p_{CT}) was used as output. AI, artificial intelligence. CT, computed tomography.

The first model was a logistic regression (LR) model with L2 regularization to prevent overfitting,[4] whereas the second model was a random forest model with 100 trees (RF-100), Gini criterion, minimum of two samples for splitting, minimum of one sample in leaves, and bootstrap.[4] The third model was a random forest model with parameters as described above, except for the limit of 10 trees and maximum depth h_max=6 (RF-10).[4] The performance of the machine-learning models was evaluated based on sensitivity, specificity, AUC, and F1-score.

Three combinations of input variables were evaluated: 1) clinical variables (mMRC, SpO_2, and FVC); 2) CXR; and 3) clinical variables (mMRC, SpO_2, FVC) and CXR.

The performance of the LR model was better when a combination of all variables (clinical variables and CXR) was used. The following metrics expressed

in terms of mean ± standard deviation and 95% Confidence Interval (CI) were considered: sensitivity, 0.85±0.08 (95% CI [0.77, 0.94]); specificity, 0.70±0.14 (95% CI [0.55, 0.85]); F1-score, 0.79±0.06 (95% CI [0.73, 0.85]); and AUC, 0.80±0.07(95% CI [0.72, 0.87]) (Table 3).

Table 3. Predictive performance of three multivariate models using three datasets.

Groups of variables	Method	Sensitivity	Specificity	F1-score	AUC
1 SpO_2, mMRC score, and FVC	LR	0.87±0.16	0.42±0.33	0.71±0.03	0.68±0.10
	RF-10	0.88±0.15	0.37±0.32	0.71±0.03	0.66±0.08
	RF-100	0.82±0.12	0.44±0.13	0.69±0.08	0.62±0.12
2 CXR	LR	0.88±0.05	0.52±0.14	0.75±0.04	0.78±0.05
	RF-10	0.91±0.08	0.41±0.18	0.73±0.04	0.73±0.06
	RF-100	0.94±0.07	0.33±0.19	0.72±0.03	0.72±0.03
3 SpO_2, mMRC score, FVC and CRX	LR	**0.85±0.08**	**0.70±0.14**	**0.79±0.06**	**0.80±0.07**
	RF-10	0.85±0.09	0.61±0.22	0.76±0.04	0.76±0.08
	RF-100	0.89±0.06	0.49±0.17	0.75±0.04	0.76±0.07

Values are presented as the mean ± standard deviation after five-fold cross validation for each test fold. Area under the receiver operating characteristic curve (AUC); Accuracy (Acc); Chest X-Ray (CRX); Forced vital capacity (FVC); Logistic Regression (LR); modified Medical Research Council dyspnea scale (mMRC); Precision (Prec); Random forest (RF).

The LR model is represented by the following function:

$$p_{CT} = \beta_1 FVC^* + \beta_2\, mMRC^* + \beta_3 S_p O_2 + \beta_4 p_{CXR0} + \beta_5 p_{CXR1} + \beta_6 p_{CXR2} + \beta_7 p_{CXR3} + \beta_8 p_{CXR4}$$

$$\beta_1 = -0.3705 \;\; \beta_2 = -2.2807 \;\; \beta_3 = -0.7450 \;\; \beta_4 = 1.1257$$

$$\beta_5 = 1.4960 \;\; \beta_6 = 1.0761 \;\; \beta_7 = 0.7328 \;\; \beta_8 = -0.7613$$

where p_{CT} is the probability of the presence of abnormalities on CT images, $FVC^* = \frac{FVC_{Resting}}{2FVC_{min}}$, $mMRC^* = \frac{mMRC}{4}$, and p_{CXR0} to p_{CXR4} are the probabilities that the CXR image has findings related to sequelae from COVID-19, obtained in each fold (0 to 4) during a 5-folds cross validation. Table 4 shows the estimates for the logistic regression function.

Table 4. Estimates of the logistic regression function.						
Variable	**Estimated regression coefficient (β)**	**Estimated Standard Error**	***p*-value**	**95% CI for regression coefficient (β)**		**Estimated odds ratios**
FVC^*	-0.3705	0.3210	0.248	-0.9990	0.2580	0.6904
$mMRC^*$	-2.2807	0.3020	<0.001	-2.8730	-1.6890	0.1022
S_pO_2	-0.7450	0.2320	0.001	-1.2010	-0.2890	0.4747
p_{CXR0}	1.1257	0.4150	0.007	0.3120	1.9400	3.0824
p_{CXR1}	1.4960	0.4160	<0.001	0.6810	2.3110	4.4638
p_{CXR2}	1.0761	0.3390	0.002	0.4120	1.7410	2.9332
p_{CXR3}	0.7328	0.3380	0.030	0.0710	1.3950	2.0809
p_{CXR4}	-0.7613	0.4580	0.096	-1.6590	0.1360	0.4671
Forced vital capacity (FVC); modified Medical Research Council dyspnea scale (mMRC); radiographic probabilities (Pcxr0 to Pcxr4).						

Dataset and normalization of clinical data

A total of 257 patients with data on the mMRC dyspnea scale, oximetry, spirometry, CRX, and chest CT were selected to predict pulmonary changes. Of the 257 patients, 128 had no significant CT changes (scores < 7). A CT score of 7 was used as the cutoff value by maximizing F1 scores and AUC (Figure 2).

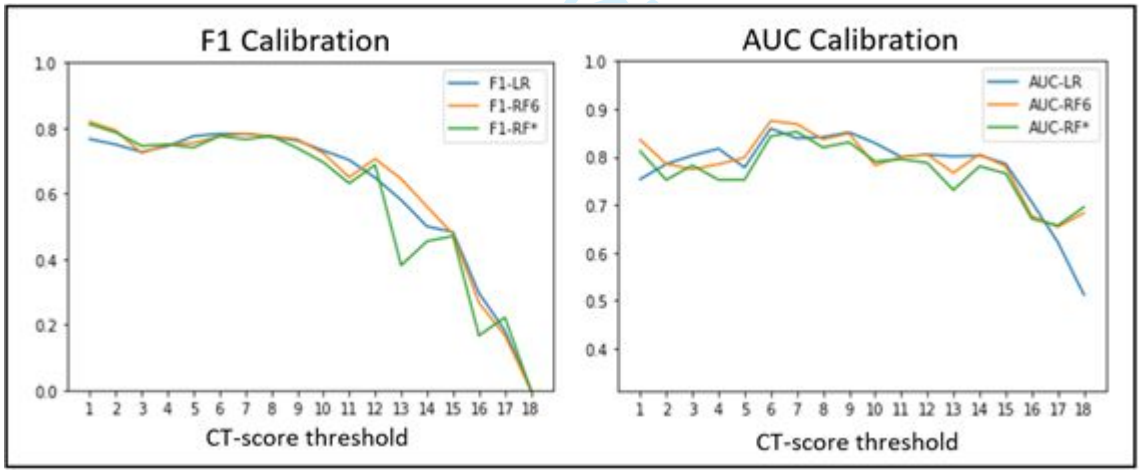
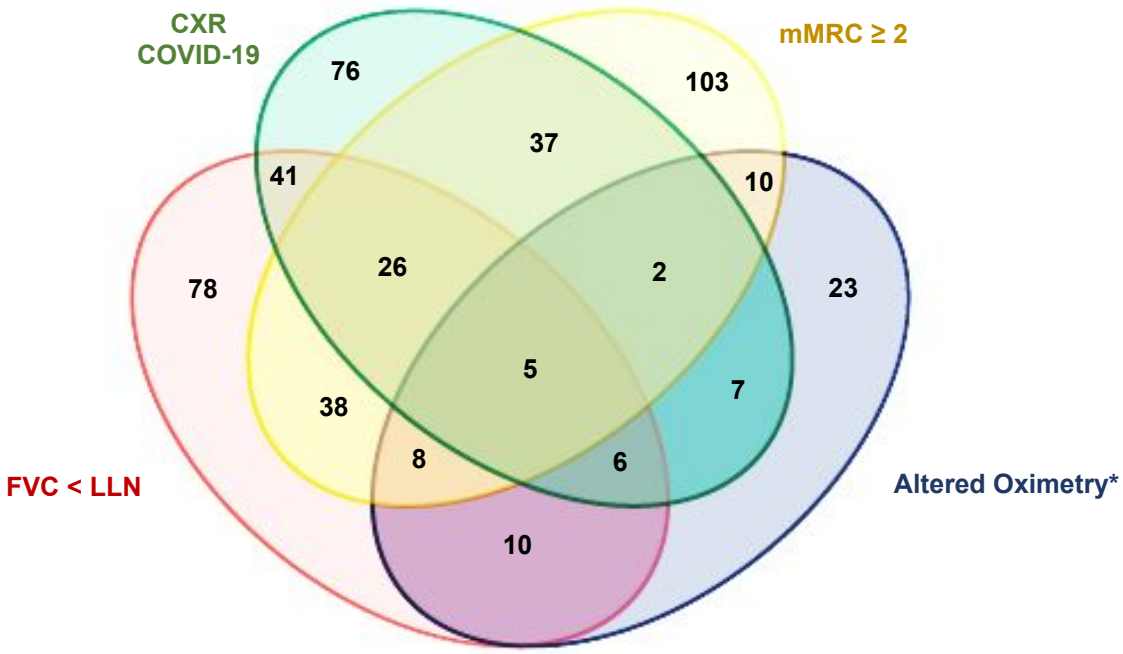
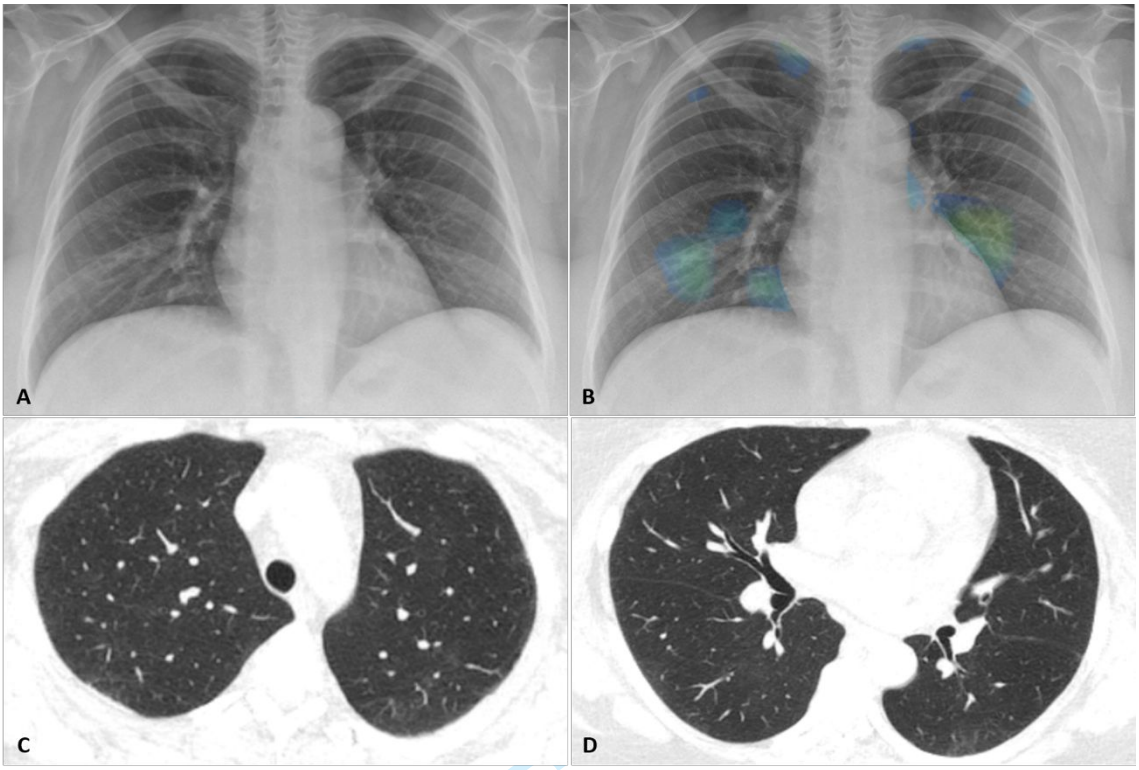


Figure 2. Computed tomography scores based on the F1-score and AUC values.

Clinical variables were normalized by dividing the mMRC values by 4 (resulting in values between 0 and 1) and the $FVC_{Resting}$ by twice the FVC_{min} (resulting in a minimum value of 0.257 and a maximum value of 0.847).



Supplemental Figure S1. Diagram showing the overlap in the changes of parameters used as pulmonary criteria to refer patients for thorax computed tomography. Values are expressed as the number of patients showing the correspondent alterations. CXR, chest X-Ray; FVC, forced vital capacity; LLN, lower limit of normal; mMRC, modified Medical Research Council dyspnea scale. *Resting SpO_2 ≤ 90% or a decrease in SpO_2 of ≥ 4% during the 1 min sit-and-stand test.



Supplemental Figure 2. Representative scan of a patient in her late 40s showing resolving ground glass abnormality after moderate COVID-19. (A) PA chest radiograph obtained 8 months after admission was considered normal by radiologists. (B) The same radiograph analyzed by the AI algorithm with heat map. Small focal abnormalities in the apical and paracardiac regions of the lungs are highlighted in green and blue. (C, D) Chest CT obtained 11 months after admission shows mild residual ground glass abnormality in the periphery of the upper lobes and left lower lobe. The patient complained of dyspnea (mMRC=3) but had normal lung function (FVC=3.81 L/91% pred) and normal oximetry (99%).

Supplemental Table S1. Demographic and clinical characteristics of the cohort of post-COVID-19 patients in this study (N=749).

Variables	Values
Age (years)	56.1 (44.4–65.1)
Male sex	399 (53.3)
BMI (kg/m^2)	30.8 (27.7–35.6) {746}
Comorbidities	
Hypertension	425 (56.7)
Smokers	285/743 (38.4)
Diabetes	261 (34.8)
COPD	55 (7.3)
Admission	
ICU	445 (59.4)
Length of ICU stay (days)	10 (6–18) {445}
IMV	304/445 (68.3)
Vital signs	
Body temperature (°C)	36.1 (35.6–36.0) {748}
Systolic blood pressure (mmHg)	124 (116–135) {743}
Diastolic blood pressure (mmHg)	77 (70–84) {743}
Heart rate (bpm)	73 (67–83) {747}
Respiratory rate (rpm)	20 (18–2) {736}
Oxygen saturation (%)	97 (95.2–98) {746}

Values are presented as median (IQR), median (IQR) {n}, n (%), or n/N (%). COPD, chronic obstructive pulmonary disease; BMI, body mass index; ICU, intensive care unit. IMV, invasive mechanical ventilation.

Supplemental Table S2. Demographic and clinical characteristics of patients with and without pulmonary involvement (N=749).

Variables	Pulmonary involvement (n=470)	No pulmonary involvement (n=279)	*p*-value
Age (years)	57.9 (45.7–65.8)	53.9 (42.5–63.7)	0.000
Male sex	228 (48.5)	171 (61.3)	0.001
BMI (kg/m^2)	31.2 (27.7–35.9) {469}	30.5 (27.6–35.2) {277}	0.111
Comorbidities			
Hypertension	287 (61.1)	138 (49.5)	0.000
Smokers	188/468 (40.2)	97/275 (35.3)	0.104
Diabetes	179 (38.1)	82 (29.4)	0.009
COPD	42 (8.9)	13 (4.7)	0.044
Admission			
ICU	317 (67.4)	128 (45.9)	0.000
Length of ICU stay (days)	11 (6–20) {317}	8 (4–14) {128}	0.000
IMV	222/317 (70)	82/128 (64.1)	0.260

Values are presented as median (IQR), median (IQR) {n}, n (%), or n/N (%). COPD, chronic obstructive pulmonary disease; BMI, body mass index; ICU, intensive care unit. IMV, invasive mechanical ventilation.

Supplemental Table S3. Demographic and clinical characteristics of COVID-19 patients with signs of pulmonary involvement (N=470).

Variables	Patients with signs of pulmonary involvement		*p*-value
	Those who underwent CT (n=348)	Those who did not undergo CT (n=122)	
Age (years)	57.8 (45.7–65.8)	58.1 (45.3–65.8)	0.490
Male sex	163 (46.8)	65 (53.3)	0.392
BMI (kg/m^2)	31.6 (28.0–36.0)	30.3 (27.0–35.9) {121}	0.041
Comorbidities			
Hypertension	215 (61.8)	72 (59)	0.469
Smokers	139/347 (40.1)	49/121 (40.5)	0.762
Diabetes	142 (40.8)	37 (30.3)	0.999
COPD	32 (9.2)	10 (8.2)	0.826
Admission			
ICU	237 (68.1)	80 (65.6)	0.999
Length of ICU stay (days)	11 (6–20) {237}	10 (4.7–19) {80}	0.913
IMV	174/237 (73.4%)	48/80 (60%)	0.034

Values are presented as median (IQR), median (IQR) {n}, n (%), or n/N (%). COPD, chronic obstructive pulmonary disease; BMI, body mass index; ICU, intensive care unit. IMV, invasive mechanical ventilation.

Supplemental Table S4. Chest computed tomography (CT) features in COVID-19 patients with CT score ≥ 7 (N=156).	
Variables	**CT changes**
CT score ≥ 7	156/328 (47.6)
Characteristics (n=156)	
Ground-glass opacities	153 (98.1)
Parenchymal bands	143 (91.7)
Reticulations	134 (85.9)
Traction bronchiectasis	92 (59)
Architectural distortion	73 (46.8)
Perilobular opacities	50 (32.1)
Bronchial wall thickening	38 (24.4)
Mosaic attenuation pattern	32 (20.5)
Consolidations	3 (1.9)
Pneumatocele	2 (1.3)
Honeycombing	-
Of the 328 patients who underwent CT scan, 47.6% had a CT score ≥ 7. Values are n/N (%) or n (%).	

Supplemental Table S5. Computed tomography changes 6 to 11 months after hospitalization due to COVID-19 (N=328).

Characteristics	Total cohort (N=328)	ICU Patients (N=222)	Ward Patients (N=106)
Ground-glass opacities	251 (76.5)	197 (86.6)	54 (51.3)
Parenchymal bands	209 (63.7)	169 (76.5)	40 (41)
Reticulations	169 (51.5)	145 (66.5)	24 (23.1)
Traction bronchiectasis	98 (29.9)	91 (44.1)	7 (7.7)
Architectural distortion	78 (23.8)	73 (35.8)	5 (6.4)
Bronchial wall thickening	89 (27.1)	60 (27.4)	29 (25.6)
Mosaic attenuation pattern	58 (17.7)	46 (20.1)	12 (11.5)
Perilobular opacities	50 (14)	47 (24.6)	3 (2.6)
Consolidation	3 (0.9)	3 (1.7)	-
Pneumatocele	2 (0.6)	2 (1.1)	-
Honeycombing	-	-	-

Values are presented as n (%).

Supplemental Table S6. Demographic and clinical characteristics of COVID-19 patients with pulmonary involvement stratified by inclusion in prediction analysis of pulmonary changes (N=328).

Variables	Patients with Pulmonary Changes		*p*-value
	Included Patients (N=257)	Excluded Patients (N=91)	
Age (years)	56.5 (45.7–64.4)	60.5 (46.9–69.9)	0.011
Male sex	113 (44)	50 (54.9)	0.068
BMI (kg/m^2)	32 (28.8–36.8)	30.6 (26.8–35.4)	0.054
Comorbidities			
Hypertension	151 (58.7)	64 (70.3)	0.060
Smokers	97/256 (37.9)	42 (46.1)	0.173
Diabetes	103 (40.1)	39 (42.9)	0.710
COPD	20 (7.8)	12 (13.2)	0.141
Admission			
ICU	179 (69.6)	58 (63.7)	0.359
Length of ICU stay (days)	12 (6–20.5) {179}	9.5 (6.2–19.7) {58}	0.209
IMV	140 (54.7)	35 (38.6)	0.010

Values are presented as median (IQR), median (IQR) {n}, n (%), or n/N (%). BMI, body mass index; COPD, chronic obstructive pulmonary disease; ICU, intensive care unit. IMV, invasive mechanical ventilation.

Supplemental References

1. Stephens K. SIIM, FISABIO, and RSNA Host Machine Learning Challenge for COVID-19 Detection and Localization. . *AXIS Imaging News* . 2021.

2. Tan M, Le Q. Efficientnet: Rethinking model scaling for convolutional neural networks. *International Conference on Machine Learning* 2019. p. 6105-14.

3. Russakovsky O, Deng J, Su H, et al. ImageNet Large Scale Visual Recognition Challenge. *International Journal of Computer Vision* 2015; **115**: 211-52.

4. Vittinghoff E, Glidden DV, Shiboski SC, et al. Regression Methods in Biostatistics: Linear, Logistic, Survival, and Repeated Measures Models. 2nd ed. New York: Springer-Verlag, 2012:1272.

TRIPOD Checklist: Prediction Model Development and Validation

Section/Topic	Item		Checklist Item	Page
Title and abstract				
Title	1	D;V	Identify the study as developing and/or validating a multivariable prediction model, the target population, and the outcome to be predicted.	1
Abstract	2	D;V	Provide a summary of objectives, study design, setting, participants, sample size, predictors, outcome, statistical analysis, results, and conclusions.	2
Introduction				
Background and objectives	3a	D;V	Explain the medical context (including whether diagnostic or prognostic) and rationale for developing or validating the multivariable prediction model, including references to existing models.	4
	3b	D;V	Specify the objectives, including whether the study describes the development or validation of the model or both.	5
Methods				
Source of data	4a	D;V	Describe the study design or source of data (e.g., randomized trial, cohort, or registry data), separately for the development and validation data sets, if applicable.	5
	4b	D;V	Specify the key study dates, including start of accrual; end of accrual; and, if applicable, end of follow-up.	5
Participants	5a	D;V	Specify key elements of the study setting (e.g., primary care, secondary care, general population) including number and location of centres.	5
	5b	D;V	Describe eligibility criteria for participants.	5
	5c	D;V	Give details of treatments received, if relevant.	-
Outcome	6a	D;V	Clearly define the outcome that is predicted by the prediction model, including how and when assessed.	7
	6b	D;V	Report any actions to blind assessment of the outcome to be predicted.	Data Supplement (Pg. 3 and 5)
Predictors	7a	D;V	Clearly define all predictors used in developing or validating the multivariable prediction model, including how and when they were measured.	6, 7
	7b	D;V	Report any actions to blind assessment of predictors for the outcome and other predictors.	Data Supplement (Pg. 3 and 5)
Sample size	8	D;V	Explain how the study size was arrived at.	5
Missing data	9	D;V	Describe how missing data were handled (e.g., complete-case analysis, single imputation, multiple imputation) with details of any imputation method.	Data Supplement, (Pg. 6)
Statistical analysis methods	10a	D	Describe how predictors were handled in the analyses.	Data Supplement (Pg. 3 and 5)
	10b	D	Specify type of model, all model-building procedures (including any predictor selection), and method for internal validation.	Data Supplement (Pg. 3 and 5)
	10c	V	For validation, describe how the predictions were calculated.	Data Supplement (Pg. 3 and 5)
	10d	D;V	Specify all measures used to assess model performance and, if relevant, to compare multiple models.	Data Supplement (Pg. 3 and 5)
	10e	V	Describe any model updating (e.g., recalibration) arising from the validation, if done.	S Data Supplement (Pg. 4)
Risk groups	11	D;V	Provide details on how risk groups were created, if done.	n.a.
Development vs. validation	12	V	For validation, identify any differences from the development data in setting, eligibility criteria, outcome, and predictors.	Data Supplement (Pg. 3 and 5)
Results				
Participants	13a	D;V	Describe the flow of participants through the study, including the number of participants with and without the outcome and, if applicable, a summary of the follow-up time. A diagram may be helpful.	8
	13b	D;V	Describe the characteristics of the participants (basic demographics, clinical features, available predictors), including the number of participants with missing data for predictors and outcome.	8
	13c	V	For validation, show a comparison with the development data of the distribution of important variables (demographics, predictors and outcome).	Data Supplement (Pg. 3 and 5)
Model development	14a	D	Specify the number of participants and outcome events in each analysis.	Data Supplement (Pg. 3 and 5)
	14b	D	If done, report the unadjusted association between each candidate predictor and outcome.	Data Supplement (Pg. 3 and 5)
Model specification	15a	D	Present the full prediction model to allow predictions for individuals (i.e., all regression coefficients, and model intercept or baseline survival at a given time point).	Data Supplement (Pg. 3 and 5)

TRIPOD Checklist: Prediction Model Development and Validation

	15b	D	Explain how to the use the prediction model.	Data Supplement (Pg. 3 and 5)
Model performance	16	D;V	Report performance measures (with CIs) for the prediction model.	Data Supplement (Pg. 5)
Model-updating	17	V	If done, report the results from any model updating (i.e., model specification, model performance).	Data Supplement (Pg.5)
Discussion				
Limitations	18	D;V	Discuss any limitations of the study (such as nonrepresentative sample, few events per predictor, missing data).	12
Interpretation	19a	V	For validation, discuss the results with reference to performance in the development data, and any other validation data.	Data Supplement (Pg.5)
	19b	D;V	Give an overall interpretation of the results, considering objectives, limitations, results from similar studies, and other relevant evidence.	10, 11, 12
Implications	20	D;V	Discuss the potential clinical use of the model and implications for future research.	10, 11, 12
Other information				
Supplementary information	21	D;V	Provide information about the availability of supplementary resources, such as study protocol, Web calculator, and data sets.	n.a.
Funding	22	D;V	Give the source of funding and the role of the funders for the present study.	14

*Items relevant only to the development of a prediction model are denoted by D, items relating solely to a validation of a prediction model are denoted by V, and items relating to both are denoted D;V. We recommend using the TRIPOD Checklist in conjunction with the TRIPOD Explanation and Elaboration document.

BMJ Open

Chronic lung lesions in COVID-19 survivors: predictive clinical model

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2021-059110.R2
Article Type:	Original research
Date Submitted by the Author:	02-Feb-2022
Complete List of Authors:	Carvalho, Carlos; Universidade de São Paulo, Instituto do Coração - Divisão de Pneumologia Chate, Rodrigo; Universidade de São Paulo Hospital das Clínicas, Instituto de Radiologia Sawamura, Marcio; Universidade de Sao Paulo Hospital das Clinicas, Instituto de Radiologia Garcia, Michelle; Universidade de São Paulo Hospital das Clínicas, Instituto do Coração - Divisão de Pneumologia Lamas, Celina ; Universidade de Sao Paulo Hospital das Clinicas, Instituto do Coração - Divisão de Pneumologia Cardenas, Diego; Universidade de São Paulo Hospital das Clínicas, Instituto do Coração - Divisão de Informática Lima, Daniel Mario; Universidade de São Paulo Hospital das Clínicas, Instituto do Coração - Divisão de Informática Scudeller, Paula; Universidade de São Paulo Hospital das Clínicas, Instituto do Coração - Divisão de Pneumologia Salge, João; Universidade de São Paulo Hospital das Clínicas, Instituto do Coração - Divisão de Pneumologia Nomura, Cesar; Universidade de São Paulo Hospital das Clínicas, Instituto de Radiologia Gutierrez, Marco; Universidade de São Paulo Hospital das Clínicas, Instituto do Coração - Divisão de Pneumologia
Primary Subject Heading:	Respiratory medicine
Secondary Subject Heading:	Global health
Keywords:	COVID-19, Chest imaging < RADIOLOGY & IMAGING, RESPIRATORY MEDICINE (see Thoracic Medicine)

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Chronic lung lesions in COVID-19 survivors: predictive clinical model

Carlos R R Carvalho (0000-0002-1618-8509)[1], Rodrigo C Chate[2], Marcio VY Sawamura[2], Michelle L Garcia[1], Celina A Lamas[1], Diego AC Cardenas[3], Daniel M Lima[3], Paula G Scudeller[1], João M Salge[1], Cesar H Nomura[2], Marco A Gutierrez[3], HCFMUSP Covid-19 Study Group*

1 Pulmonary Division, Heart Institute (InCor), Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo (HCFMUSP), Sao Paulo, SP, Brazil.

2 Radiology Institute (InRad), Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo (HCFMUSP), Sao Paulo, SP, Brazil.

3 Informatics Division, Heart Institute (InCor), Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo (HCFMUSP), Sao Paulo, SP, Brazil.

Correspondence to: Dr Carlos RR Carvalho (ORCID: 0000-0002-1618-8509), Pulmonary Division, Heart Institute (InCor), Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo (HCFMUSP), Av. Dr Eneas Carvalho de Aguiar, 44, Cerqueira Cesar, São Paulo, SP, 05403-900. Email: carlos.carvalho@hc.fm.usp.br. Fone: +55 11 26614505.

Word Count: 3550

Abstract

Objective This study aimed to propose a simple, accessible, and low-cost predictive clinical model to detect lung lesions due to COVID-19 infection.

Design, settings and participants: This prospective cohort study included COVID-19 survivors hospitalised between March 30, 2020 and August 31, 2020 followed-up after six months of discharge from a tertiary hospital in Sao Paulo, Brazil. There were 749 eligible RT-PCR-confirmed SARS-CoV-2 infected patients aged ≥18 years (median [IQR] age, 56 [44.4–65.1] years; 53% male). 257 patients had complete data and were included for the prediction analysis of pulmonary changes.

Outcome Measures: Anthropometric data and pulmonary function were assessed using the modified Medical Research Council (mMRC) dyspnoea scale, oximetry (SpO_2), spirometry (forced vital capacity [FVC]), and chest X-ray (CXR) during an in-person consultation. Patients with abnormalities in at least one of these parameters underwent chest computed tomography (CT). The median interval between hospital admission and consultation was 7.1 [6.7–8.5] months, and that between the first in-person consultation and chest CT was 45±33 days. mMRC scale, SpO_2, FVC, and CXR findings were used to build a machine learning model for lung lesion detection on CT.

Results There were 470 patients (63%) that had at least one sign of pulmonary involvement and were eligible for CT. 48% of them had significant pulmonary abnormalities, including ground-glass opacities, parenchymal bands, reticulation, traction bronchiectasis, and architectural distortion. The machine learning model accurately detected pulmonary lesions by the joint data of CXR, mMRC scale, SpO_2, and FVC (sensitivity, 0.85±0.08; specificity, 0.70±0.06; F1-score, 0.79±0.06; and AUC, 0.80±0.07).

Conclusion A predictive clinical model based on CXR, mMRC, oximetry, and spirometry data can accurately screen patients with lung lesions after SARS-CoV-2 infection. Given that these examinations are highly accessible and low cost, this protocol can be automated and implemented in different countries for early detection of COVID-19 sequelae.

Strengths and limitations of this study

- This study conducted a broad assessment, embracing an in-person clinical, functional, and radiological pulmonary examinations of a large cohort of COVID-19 patients.

- The sample size used for artificial intelligence evaluation was sufficient to provide a robust prediction equation.

- Although the study was conducted in a single centre, the cohort population was heterogeneous and hailed from all districts of the metropolitan region of Sao Paulo (with approximately 21 million inhabitants).

- Although there were some missing patient data and data lost to follow-up, in general they were from patients that had less severe disease.

INTRODUCTION

The coronavirus disease 2019 (COVID-19) caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) emerged in December 2019 and had since spread globally.[1] This multisystemic viral disease promotes endothelial and microvascular damage and immune system dysregulation, leading to hyperinflammatory and hypercoagulable states.[2 3] Several organs can be affected during the acute phase of COVID-19. In particular, pulmonary complications are considered life-threatening owing to the risk of progression to respiratory failure.[4 5]

COVID-19 symptoms can persist for more than 12 weeks after acute infection, characterizing long COVID.[1] The clinical complains of dyspnoea, fatigue, cough, chest pain, depression, cognitive disorders, headache, palpitations, myalgia, and arthralgia are the most reported in long COVID.[6-9] In addition to symptoms, some studies have shown that radiological abnormalities are also frequent in the follow-up of patients after the acute phase. One study performed chest computed tomography (CT) in 171 patients 4 months after hospital discharge and showed abnormalities in 75.5% of the patients who required invasive mechanical ventilation (IMV).[10] “Fibrotic-like changes” were observed in 19.3% of the total cohort and in 38.8% of patients with acute respiratory distress syndrome.[9] IMV can predict pulmonary sequelae, which reduce functional capacity and the health-related quality of life.[6 11 12] The National Institute for Health and Care Excellence (NICE), has reported that some examinations can guide the diagnosis and management of post-COVID-19 syndrome,[1] including oximetry, spirometry, chest X-ray (CXR), ultrasonography, modified Medical Research Council (mMRC) dyspnoea scale, and chest CT. The latter examination is the gold standard for the diagnosis of chronic lung lesions due to COVID-19 and characterization of “fibrotic-like” lung lesions.[1 10]

The World Health Organization reported more than 265 million confirmed COVID-19 cases worldwide, with approximately 5 million deaths, and 260 million patients recovered as of December 2021.[13] The large number of recovered individuals experiencing long-term COVID-19 symptoms, such as fatigue, weakness, and dyspnoea, has drawn the attention of researches,[14 15] as they are

expected to impose a significant health and economic burden.[14] In early 2021, the United Kingdom National Institute for Health Research invested £18.5 million to fund studies on long COVID.[16] The lack of knowledge and medical training for treating post-COVID symptoms also represents a significant public health challenge.[14] Thus, health care systems will have to reorganize themselves to address this issue, requiring the reallocation of resources and training of multidisciplinary teams and the development of new approaches.[14]

In this context, the wide availability of CXR and CT scanners has enabled the development of deep learning (DL) artificial intelligence-based algorithms for the automated diagnosis and prognosis of COVID-19.[17-19] For example, Castiglioni et al. [17] proposed a DL model for diagnosing COVID-19 with high sensitivity and specificity using radiography findings, whereas Wang et al. [18] developed a DL model (DenseNet) to classify CT images as positive or negative for COVID-19.

Although these studies presented promising results, they focused on images of patients in the acute phase of COVID-19. However, as the pandemic is still ongoing with limited knowledge on long COVID-19 consequences,[20] a more comprehensive protocol for screening COVID-19 patients and assessing the risk of chronic pulmonary changes in recovered patients has not been validated to date. Thus, this study aimed to propose a predictive clinical model to detect the presence of radiologic chronic lung lesions due to SARS-CoV-2 infections based on the results of simple and accessible examinations, such as the mMRC dyspnoea scale, oximetry, spirometry, and CXR.

METHODS

Study design and eligibility

This prospective cohort study detected lung lesions in adult patients (≥ 18 years) with RT-PCR-confirmed SARS-CoV-2 infection admitted to the ward or intensive care unit (ICU) of the Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo (HCFMUSP), Sao Paulo, Brazil, from March 30 to August 31st, 2020. The RT-PCR-confirmed SARS-CoV-2 infection was obtained at hospital admission day. We considered only the first admission of each patient on the HCFMUSP. The protocols used in this study were described previously.[21]

All research procedures were approved by the Research Ethics Committee of our institution (Process No. 31942020.0.000.0068).

The patients were invited to participate in the study six months after admission, and a face-to-face consultation was scheduled. At this point, all patients were already discharged. Clinical, radiological, and laboratory evaluations were performed at face-to-face consultations after the patients gave written informed consent. Clinical data (comorbidities, cardiorespiratory symptoms, and smoking history), including the length of ICU stay and the need for IMV, were retrospectively collected from the electronic medical records of HCFMUSP. All data were stored in a structured form developed using REDCap software (https://www.redcapbrasil.com.br/).

General evaluation

Patients who agreed to participate in the study signed an informed consent form and underwent a face-to-face consultation during the collection of anthropometric data and a pulmonary assessment, with an emphasis on respiratory symptoms. Dyspnoea was assessed using the mMRC scale.[21] Oxygen saturation (SpO_2) at rest and after physical exertion (1-min sit and stand test) was measured by pulse oximetry.[21 22] Spirometry was performed according to criteria established by ATS/ERS Task Force.[23] Actual spirometry results were compared with predicted values, according to Pereira et al. [24].

At the same face-to-face consultation described above, the same patients underwent a posteroanterior and lateral CXR according to standard guidelines. The results of these examinations were evaluated blindly and independently by two chest radiologists (MVYS and RCC, have 7 and 16 years of experience in thoracic radiology, respectively) working on dedicated workstations. The radiographs were scored as 0 (results were normal or not related to COVID-19 [including cardiomegaly and pulmonary nodules, for instance]) or 1 (findings which could be related to COVID-19 [including bilateral linear and/or reticular opacities, especially peripheral opacities]). Disagreements were resolved by consensus. The agreement rate was 75%.

After the consensus classification performed by the radiologists (described above), the dataset with classified CXR were used to train and validate a DL algorithm developed to predict the probability that the CXR had findings related to sequelae of COVID-19. The DL algorithm is based on an EfficientNetB7 architecture[25] and a five-fold cross-validation strategy was adopted to train and validate the model, leading to an average area under the curve (AUC) of 0.89 (Supplemental Methods [Table 2]).

Chest CT

Patients who meet at least one the following criteria during the general evaluation were enrolled to undergo CT: (a) mMRC ≥ 2; (b) resting SpO_2 ≤ 90% and/or a decrease in SpO_2 of ≥4% during the 1-min sit and stand test; (c) opacities likely related to COVID-19 on CXR; and (d) FVC < lower limit of normal (LLN). The mean interval between CXR and chest CT was 45 ± 33 days.

The CT protocol used in this study was described previously.[21] CT findings consistent with COVID-19, including ground-glass and peripheral opacities, consolidations, parenchymal bands, reticulations, traction bronchiectasis, architectural distortions, honeycombing, bronchial wall thickening, mosaic attenuation, and pleural effusion, were categorized according to the criteria of the Fleischner Society.[26] The extent of lung involvement was quantified according to Francone et al. [27] by assigning the following scores to each pulmonary lobe: 0, none; 1, <5%; 2, 5-25%; 3, 26-50%; 4, 51-75%; and 5, >75%. The total score varied from 0 to 25 and was calculated by summing the scores of the five lobes. [25] Categorization of the CT features and score assignment were blindly and independently performed by the same two thoracic radiologists who evaluated the CXR (MVYS and RCC). Any disagreements were resolved by consensus.

A score ≥7 was used as the cut off value for significant CT changes after model calibration. The equations used to determine these scores are described in the Supplemental Methods.

Machine learning (ML) model

A Machine Learning (ML) model based on a Logistic Regression (LR) with L2 regularization to prevent overfitting[28] was adopted to detect the presence of COVID-19-related chronic lung lesions. The L1 regularization was not included due to the variable selection by statistical significance that removed irrelevant and correlated attributes. In this ML model, the results of the mMRC scale, oximetry, spirometry, and DL-based classification of 257 CXR images were used as input data, and the presence of pulmonary lesions was used as output data (Figure 1). The performance of the model was evaluated by the metrics sensitivity, specificity, AUC, and F1-score after a five-fold cross validation. (Supplemental Methods)

Statistical analysis

Continuous variables are expressed as the mean and standard deviations or median and interquartile range. Normality of the variables was assessed by D'agostino-Pearson test. Normally and non-normally distributed continuous variables were compared using the Student's *t*-test and Mann-Whitney U test, respectively. Categorical variables are presented as counts and percentages and compared using the chi-square test. (Excel 2016; Python 3.8.11; extension packages: Pandas 1.0.1; Numpy 1.19.5; Scipy 1.5.4; Scikit-Learn 0.24.0).

The performance of the DL models was assessed by the area under the receiver operating characteristic (AUC) curve. The performance of the ML model was determined based on sensitivity, specificity, F1-score and AUC values (Supplemental Methods).

Patient and public involvement

Patients or the public were not involved in the design, conduct, reporting or dissemination plans of this research.

RESULTS

Of 3,753 COVID-19 enrolled patients, 1,957 were eligible for the study and 749 were included in the final analysis (445 [59%] and 304 [41%] patients were admitted to the ICU and ward, respectively). Additional information on the inclusion and exclusion criteria is shown in Figure 2.

Demographic characteristics of the cohort are shown in Supplemental Table S1. The median age was 56 years, with a predominance of overweight individuals, and 53% were male. Additionally, 59.4% of the patients were admitted to the ICU, and 68.5% of them were on IMV during the study period. The vital signs of most patients were within normal limits during the hospitalisation period (Supplemental Table S1).

The median interval between hospital admission and consultation was 7.1 (IQR [6.7–8.5]) months, and the lower and upper limits of the median were 5.4 and 12.9 months, respectively. Of the 749 patients, 470 (63%) had at least one sign of pulmonary involvement (Table 1). Supplemental Figure S1 illustrates the simultaneous presence of two or more criteria for pulmonary involvement.

Table 1. Pulmonary function of patients with signs of pulmonary involvement (N=749).

Variables	Patients with signs of pulmonary involvement (N=749)
mMRC ≥ 2	229/742 (30.9)
Altered Oximetry*	71/675 (10.5)
CXR (score 1)	200/629 (31.8)
FVC < LLN	212/642 (33)

Values are presented as n/N (%). CXR, chest X-ray; FVC, forced vital capacity; mMRC, modified Medical Research Council dyspnoea scale; LLN, lower limit of normal. *Resting SpO_2 ≤90% or a decrease in SpO_2 of ≥4% during the 1-min sit and stand test.

The demographic and clinical characteristics of patients stratified by the presence of pulmonary involvement are described in Supplemental Table S2. Patients with pulmonary involvement were older and predominantly female, have more comorbidities, and a higher rate of ICU admission than those without (Supplemental Table S2). In patients with pulmonary involvement, 348 underwent

CT (68%) (Figure 2). The demographic and clinical characteristics were similar between those that underwent or did not undergo the CT (Supplemental Table S3).

CT scores were obtained from 328 (94%) patients. Scores were not determined in 20 patients, who were excluded because of low CT scan quality or had motion artefacts. Chest CT analysis showed that 47.6% of the patients had a score ≥7, and the most common features were ground-glass opacities, parenchymal bands, reticulation, traction bronchiectasis, and architectural distortions (Supplemental Table S4). In this group, 86.5% and 13.5% were admitted to the ICU and ward, respectively. Among the patients with normal CT findings (score = 0), 36.4% and 63.6% were admitted to the ICU and ward, respectively. The frequency of CT changes is shown in Supplemental Table S5. That frequency of "fibrotic-like" lesions, including traction bronchiectasis and architectural distortion, was significantly higher in the group admitted to the ICU in the acute phase of the disease. Long-term CT features in patients with moderate and critical COVID-19 are shown in Figure 3 and Supplemental Figure S2, respectively.

Of the 348 patients with CT data, 257 had data on mMRC, oximetry, spirometry, X-ray, and chest CT and were selected for the prediction analysis of pulmonary changes. Among the 91 patients excluded for the prediction analysis, 61 had incomplete data of all four tests (mMRC, oximetry, spirometry, CXR and CT) and 30 showed radiographic signs not related to COVID-19 (Supplemental Table S6).

Three data groups were considered for the prediction analysis of pulmonary changes: (1) clinical data (oximetry [SpO_2], mMRC dyspnoea scores, and spirometry [FVC]), (2) CXR, and (3) all variables (oximetry [SpO_2], mMRC dyspnoea scores, spirometry [FVC], and CXR). The performance of the predictive model was higher using the combination of all variables (clinical variables and CXR), and the following metrics expressed in terms of mean ± standard deviation and 95% Confidence Interval (CI) were considered: sensitivity, 0.85±0.08 (95% CI [0.77, 0.94]); specificity, 0.70±0.14 (95% CI [0.55, 0.85]); F1-score, 0.79±0.06 (95% CI [0.73, 0.85]); and AUC, 0.80±0.07(95% CI [0.72, 0.87]) (Table 2).

Table 2. Performance of the predictive model using three combinations of variables (N=257).				
Groups of variables	**Sensitivity**	**Specificity**	**F1-score**	**AUC**
1 SpO_2, mMRC score, and FVC	0.87±0.16	0.42±0.33	0.71±0.03	0.68±0.10
2 CXR	0.88±0.05	0.52±0.14	0.75±0.04	0.78±0.05
3 SpO_2, mMRC score, FVC, and CXR	0.85±0.08	0.70±0.14	0.79±0.06	0.80±0.07
Values are presented as means ± standard deviations after five-fold cross validation for each test fold. CXR, chest X-Ray; FVC, forced vital capacity; mMRC, Modified Medical Research Council dyspnoea scale.				

The machine learning predictive model is represented by the following function:

$$p_{CT} = \sigma(\beta_1 FVC^* + \beta_2\, mMRC^* + \beta_3 S_pO_2 + \beta_4 p_{CXR0} + \beta_5 p_{CXR1} + \beta_6 p_{CXR2} + \beta_7 p_{CXR3} + \beta_8 p_{CXR4})$$

$$\beta_1 = -0.3705 \;\; \beta_2 = -2.2807 \;\; \beta_3 = -0.745 \;\; \beta_4 = 1.1257$$

$$\beta_5 = 1.4960 \;\; \beta_6 = 1.0761 \;\; \beta_7 = 0.7328 \;\; \beta_8 = -0.7613$$

Where p_{CT} is the probability of the presence of abnormalities on CT images, σ is the sigmoid function to restrict p_{CT} between 0 and 1, $FVC^* = \frac{FVC_{Resting}}{2FVC_{min}}$, $mMRC^* = \frac{mMRC}{4}$, and p_{CXR0} to p_{CXR4} are the probabilities that the CXR image has findings related to sequelae from COVID-19, obtained in each fold (0 to 4) during a 5-folds cross validation. (Supplemental Methods)

Therefore, based in these observations, we propose in a flowchart a suggestion for lung lesion case-finding in COVID-19 survivors (Figure 4).

DISCUSSION

Few studies have assessed the pulmonary abnormalities in COVID-19 survivors after six months of hospital discharge. However, some of these patients

have developed long-term pulmonary complications after the acute phase of the disease.[6 29-33] This study evaluated 749 COVID-19 patients who received supplemental oxygen or ventilatory support in the ward or ICU and survived. They underwent an in-person comprehensive clinical, functional, and radiological assessments, which were more extensive than those performed in previous studies,[6 30 31 33-35] conferring reliability to our results.

In the first months after recovery, the most common CT findings in COVID-19 hospitalised patients included ground-glass opacities, parenchymal bands, reticulation, mosaic attenuation pattern, and "fibrotic-like" abnormalities, including traction bronchiectasis and architectural distortions.[36 37] These findings were detected in 76.5% of our cohort, and severe and extensive changes were noted in approximately 50% of the cases. The CT abnormalities were more prevalent in older critical patients and individuals with more comorbidities, which is consistent with previous studies.[32 38] These results indicate the high prevalence of chronic lung lesions and sequelae in post-COVID patients worldwide.

Therefore, the need to identify severe pulmonary complications due to COVID-19, including fibrosis,[1] and the large number of COVID-19 survivors, prompted us to develop a predictive clinical model to screen patients admitted to a tertiary hospital, which could be able to reduce costs and radiation exposure. During the first six months of the pandemic in Sao Paulo, Brazil, all hospital beds at HCFMUSP (300 in the ICU and 400 in the ward) were made available to COVID-19 patients.[12] Patients were treated free of charge in our hospital owing to a universal health system, and there is a constant search for better and cost-effective protocols to improve workflow.[12]

Dyspnoea scales, CXR, oximetry, and spirometry are commonly used to evaluate COVID-19 symptoms.[2] A Norwegian study evaluated a cohort of 100 patients three months after admission to a hospital and reported that 19% had dyspnoea (mMRC score>1) and 10% presented altered FVC and normal oxygen saturation levels, suggesting the lower sensitivity of pulse oximetry.[39] In 113 patients evaluated 4 months after COVID-19 diagnosis in Switzerland, FVC and oxygen saturation levels were lower in patients who had a severe disease than in those with a moderate disease, although the mean values remained within the

limits of normality.[35] In addition, a previous study has suggested that cough, lymphocytosis and the lung volume could indicate lung lesions in COVID-19-recovered patients.[34]

Ground-glass and reticular opacities can be detected by CXR, although this method is less sensitive than CT.[40] On the other hand, CXR is readily available in the primary care setting and has a lower cost and radiation exposure than CT.[40 41] Radiographs were separately scored by an automated DL-based image analysis tool and chest radiology specialists, and there was a high level of consensus between these scores (AUC = 0.89). In the Brazilian public health system, the cost of a CT scan is approximately 15 times higher than that of a CXR.[41] According to the American College of Radiology and the Radiological Society of North American, the radiation doses of a standard chest CT and CXR are 6.1 mSv and 0.1 mSv, respectively; this underscores the advantage of CXR in reducing the exposure of COVID-19 patients to radiation, especially those who have already performed serial imaging exams in the acute phase of the disease.[42]

Nevertheless, none of these examinations alone accurately predicted pulmonary complications. The performance of our model corroborates this finding since the information provided by each clinical examination alone did not accurately diagnose the pulmonary changes detected on CT. In contrast, clinical and radiographic data were complementary and increased the performance of the ML model. Cross-validation also increased the robustness of the results. These results indicate that four examinations (oximetry, mMRC dyspnoea scale, spirometry, and CXR) should be jointly conducted to screen patients at risk of developing chronic lung lesions due to COVID-19 and achieve a diagnostic performance similar to that of CT (sensitivity, 0.85±0.08; specificity, 0.70±0.14; F1-score, 0.79±0.06; and AUC, 0.80±0.07). Analysis of these metrics indicates that this predictive clinical method can better identify the true positives than true negatives. In addition, the F1-score takes into account both false-positive and false-negative results and measures the accuracy of the method in the dataset.

The WHO has highlighted the importance of establishing screening protocols with a favourable cost-effectiveness ratio for patients affected by different pathologies.[43] The identification of COVID-19 lung lesions will allow the

accurate referral of patients to specialists for further investigation and treatment. As the COVID-19 sequelae can progress to increasing intensity of symptoms and risk of disability, this approach can improve the quality and length of life of patients, since medical interventions can be performed as early as possible.

We already have an initiative to implement this protocol in Brazil. The project will start in the state of Sao Paulo, in partnership with the State of Sao Paulo Health Department, where the HCFMUSP is located. We will start to apply this screening protocol in the central area of the city of Sao Paulo, with approximately 430.000 inhabitants, according to the flowchart suggested for lung lesion case-finding in COVID-19 survivors (Figure 4). Firstly, exams will be performed in the following order, starting from the simplest and most accessible ones: oximetry/mMRC, spirometry and CXR. At the moment the patient shows alterations in any of these four exams, the patient will be enrolled directly for further investigation in a specialised care centre to perform CT and/or other specific exams. We expect that over time, this can lead to a significant reduction in morbidity and mortality due to COVID-19 lung sequelae, relieving the burden on the health care system, reducing expenses of imaging exams and accelerating the medical interventions.

This study has some limitations. First, there was variability in the interval between the execution of CXR and CT. Notwithstanding this variation, which might contribute to lung recovery, our protocol screened a large number of patients with pulmonary lesions, demonstrating the persistence of these manifestations secondary to COVID-19 and reducing sampling bias. Second, the single-centre nature of the study limits the generalizability of our results. However, a previous study showed that the population of patients admitted to HCFMUSP—a tertiary reference hospital for the treatment of COVID-19 in Brazil—was heterogeneous and hailed from all districts of the metropolitan region of Sao Paulo (with approximately 21 million inhabitants).[12] Third, we were unable to contact some patients because of inconsistencies in telephone numbers and addresses. Thus, these subjects were not included in the protocol, although public death registry data showed that they were alive. Fourth, this screening protocol was developed based on respiratory complaints, which are considered

risk factors for the development of chronic lung complications. However, other COVID-19 symptoms were not analysed in this study.

The breadth of our results allowed us to propose a simple, accessible, and low-cost clinical predictive model to screen patients at risk of developing chronic lung lesions due to COVID-19. The low cost and easy accessibility to these examinations facilitate the implementation of the proposed protocol in developing countries. In addition, it may contribute to early and effective determination of the treatment course, thus reducing radiation exposure and the conduct of costly imaging examinations. The use of artificial intelligence facilitated the large-scale assessment of radiographs and their association with clinical variables, demonstrating that artificial intelligence models can be used to automate diagnosis, especially in severe patients.

Collaborators: *Members of the HCFMUSP Covid-19 Study Group: Adriana L Araújo, Aluisio C Segurado, Amanda C Montal, Anna Miethke-Morais, Anna S Levin, Beatriz Perondi, Bruno F Guedes, Carolina Carmo, Carolina S Lázari, Cassiano C Antonio, Clarice Tanaka, Claudia C Leite, Cristiano Gomes, Edivaldo M Utiyama, Emmanuel A Burdmann, Eloisa Bonfá, Esper G Kallas, Ester Sabino, Euripedes C Miguel, Fabio R Pinna, Fabiane Y O Kawano, Geraldo F Busatto, Giovanni G Cerri, Guilherme Fonseca, Heraldo P Souza, Izabel Marcilio, Izabel C Rios, Jorge Hallak, José Eduardo Krieger, Juliana C Ferreira, Julio F M Marchini, Larissa S Oliveira, Leila Harima, Linamara R Batisttella, Luis Yu, Luiz Henrique M Castro, Marcelo C Rocha , Marcello M C Magri, Marcio Mancini, Maria Amélia de Jesus, Maria Cassia J M Corrêa, Maria Cristina P B Francisco, Maria Elizabeth Rossi, Marjorie F Silva, Marta Imamura, Maura S Oliveira, Nelson Gouveia, Orestes V Forlenza, Paulo A Lotufo, Ricardo F Bento, Ricardo Nitrini, Rodolfo F Damiano, Roger Chammas, Rossana P Francisco, Solange R G Fusco, Tarcisio E P Barros-Filho, Thais Mauad, Thaís Guimarães, Thiago Avelino-Silva and Wilson J Filho.

Funding: This study was funded by the Sao Paulo Research Foundation (grant number - 2020/07200-9).

Contributors: CRRC: conceptualisation, data curation, formal analysis, funding acquisition, investigation, methodology, project administration, resources, supervision, validation, visualisation, writing – original draft, and writing – review & editing. RCC: data curation, formal analysis, investigation, methodology, validation, visualisation, writing – original draft, and writing – review & editing. MVYS: data curation, formal analysis, investigation, methodology, validation, visualisation, writing – original draft, and writing – review & editing. MLG: conceptualisation, data curation, formal analysis, investigation, methodology, project administration, supervision, validation, visualisation, writing – original draft, and writing – review & editing. CAL: data curation, formal analysis, investigation, writing – original draft, and writing – review & editing. DACC: data curation, formal analysis, methodology, software, writing – original draft, and writing – review & editing. DML: data curation, formal analysis, methodology, software, writing – original draft, and writing – review & editing. PGS: conceptualisation, project administration, supervision, validation, visualisation and writing – review & editing. JMS: methodology, validation, visualisation and writing – review & editing. CHN: methodology, validation, visualisation and writing – review & editing. MAG: data curation, formal analysis, funding acquisition, methodology, software, supervision, validation, visualisation, writing – original draft, and writing – review & editing. HCFMUSP Covid-19 Study Group: contributed to the implementation of the study and data collection. All authors critically reviewed and approved the final version.

Competing interests' statement: None declared.

Data Sharing: The study protocol was previously described by Busatto et al. [21] and was registered at the "Brazilian Registry of Clinical Trials" (https://ensaiosclinicos.gov.br/). The raw data are not publicly available because follow-up studies will be carried out. However, data are available from the corresponding author upon request and authorization from the institution. Data on demographics, hospitalisation, and outcomes are available in the COVID-19 Data Sharing/BR repository and are freely available for download[44].

Ethical approval: The study was approved by the Research Ethics Committee of the HCFMUSP (approval number 31942020.0.000.0068).

REFERENCES

1. Sisó-Almirall A, Brito-Zerón P, Conangla Ferrín L, et al. Long Covid-19: Proposed Primary Care Clinical Guidelines for Diagnosis and Disease Management. *Int J Environ Res Public Health* 2021;18(8) doi: 10.3390/ijerph18084350 [published Online First: 2021/04/20]
2. Nalbandian A, Sehgal K, Gupta A, et al. Post-acute COVID-19 syndrome. *Nat Med* 2021;27(4):601-15. doi: 10.1038/s41591-021-01283-z [published Online First: 2021/03/22]
3. Mauad T, Duarte-Neto AN, da Silva LFF, et al. Tracking the time course of pathological patterns of lung injury in severe COVID-19. *Respir Res* 2021;22(1):32. doi: 10.1186/s12931-021-01628-9 [published Online First: 20210129]
4. Tanni SE, Fabro AT, de Albuquerque A, et al. Pulmonary fibrosis secondary to COVID-19: a narrative review. *Expert Rev Respir Med* 2021;15(6):791-803. doi: 10.1080/17476348.2021.1916472 [published Online First: 2021/04/27]
5. Macedo BR, Garcia MVF, Garcia ML, et al. Implementation of Tele-ICU during the COVID-19 pandemic. *J Bras Pneumol* 2021;47(2):e20200545. doi: 10.36416/1806-3756/e20200545 [published Online First: 20210430]
6. Huang C, Huang L, Wang Y, et al. 6-month consequences of COVID-19 in patients discharged from hospital: a cohort study. *Lancet* 2021;397(10270):220-32. doi: 10.1016/S0140-6736(20)32656-8 [published Online First: 2021/01/08]
7. Lopez-Leon S, Wegman-Ostrosky T, Perelman C, et al. More Than 50 Long-Term Effects of COVID-19: A Systematic Review and Meta-Analysis. *Res Sq* 2021 doi: 10.21203/rs.3.rs-266574/v1 [published Online First: 2021/03/01]
8. Fernandes PMP, Mariani AW. Life post-COVID-19: symptoms and chronic complications. *Sao Paulo Med J* 2021;139(1):1-2. doi: 10.1590/1516-3180.2021.139104022021
9. Morin L, Savale L, Pham T, et al. Four-Month Clinical Status of a Cohort of Patients After Hospitalization for COVID-19. *JAMA* 2021;325(15):1525-34. doi: 10.1001/jama.2021.3331
10. Raghu G, Collard HR, Egan JJ, et al. An official ATS/ERS/JRS/ALAT statement: idiopathic pulmonary fibrosis: evidence-based guidelines for diagnosis and management. *Am J Respir Crit Care Med* 2011;183(6):788-824. doi: 10.1164/rccm.2009-040GL
11. Carfì A, Bernabei R, Landi F, et al. Persistent Symptoms in Patients After Acute COVID-19. *JAMA* 2020;324(6):603-05. doi: 10.1001/jama.2020.12603
12. Ferreira JC, Ho YL, Besen BAMP, et al. Protective ventilation and outcomes of critically ill patients with COVID-19: a cohort study. *Ann Intensive Care* 2021;11(1):92. doi: 10.1186/s13613-021-00882-w [published Online First: 2021/06/07]
13. WHO. WHO Coronavirus Disease (COVID-19) Dashboard 2021 [cited 2021 October 13]. Available from: https://covid19.who.int/ accessed October 13 2021
14. Wade DT. Rehabilitation after COVID-19: an evidence-based approach. *Clin Med (Lond)* 2020;20(4):359-65. doi: 10.7861/clinmed.2020-0353 [published Online First: 2020/06/09]
15. Godoy CG, Silva ECGE, Oliveira DB, et al. Protocol for Functional Assessment of Adults and Older Adults after Hospitalization for COVID-19. *Clinics (Sao Paulo)* 2021;76:e3030. doi: 10.6061/clinics/2021/e3030 [published Online First: 20210614]
16. Subbaraman N. US health agency will invest $1 billion to investigate 'long COVID'. *Nature* 2021;591(7850):356. doi: 10.1038/d41586-021-00586-y
17. Castiglioni I, Ippolito D, Interlenghi M, et al. Machine learning applied on chest x-ray can aid in the diagnosis of COVID-19: a first experience from Lombardy, Italy. *Eur Radiol Exp* 2021;5(1):7. doi: 10.1186/s41747-020-00203-z [published Online First: 20210202]

18. Wang S, Zha Y, Li W, et al. A fully automatic deep learning system for COVID-19 diagnostic and prognostic analysis. *Eur Respir J* 2020;56(2) doi: 10.1183/13993003.00775-2020 [published Online First: 20200806]
19. Ferreira Junior JR, Cardona Cardenas DA, Moreno RA, et al. Novel Chest Radiographic Biomarkers for COVID-19 Using Radiomic Features Associated with Diagnostics and Outcomes. *J Digit Imaging* 2021;34(2):297-307. doi: 10.1007/s10278-021-00421-w [published Online First: 20210218]
20. Greenhalgh T, Knight M, A'Court C, et al. Management of post-acute covid-19 in primary care. *BMJ* 2020;370:m3026. doi: 10.1136/bmj.m3026 [published Online First: 2020/08/11]
21. Busatto GF, de Araújo AL, Duarte AJDS, et al. Post-acute sequelae of SARS-CoV-2 infection (PASC): a protocol for a multidisciplinary prospective observational evaluation of a cohort of patients surviving hospitalisation in Sao Paulo, Brazil. *BMJ Open* 2021;11(6):e051706. doi: 10.1136/bmjopen-2021-051706 [published Online First: 2021/06/30]
22. van den Borst B, Peters JB, Brink M, et al. Comprehensive health assessment three months after recovery from acute COVID-19. *Clin Infect Dis* 2020 doi: 10.1093/cid/ciaa1750 [published Online First: 2020/11/21]
23. Miller MR, Hankinson J, Brusasco V, et al. Standardisation of spirometry. *Eur Respir J* 2005;26(2):319-38. doi: 10.1183/09031936.05.00034805
24. Pereira CA, Sato T, Rodrigues SC. New reference values for forced spirometry in white adults in Brazil. *J Bras Pneumol* 2007;33(4):397-406. doi: 10.1590/s1806-37132007000400008
25. Tan M, Le Q. Efficientnet: Rethinking model scaling for convolutional neural networks. *International Conference on Machine Learning* 2019:6105-14.
26. Hansell DM, Bankier AA, MacMahon H, et al. Fleischner Society: glossary of terms for thoracic imaging. *Radiology* 2008;246(3):697-722. doi: 10.1148/radiol.2462070712 [published Online First: 20080114]
27. Francone M, Iafrate F, Masci GM, et al. Chest CT score in COVID-19 patients: correlation with disease severity and short-term prognosis. *Eur Radiol* 2020;30(12):6808-17. doi: 10.1007/s00330-020-07033-y [published Online First: 2020/07/04]
28. Vittinghoff E, Glidden DV, Shiboski SC, et al. Regression Methods in Biostatistics: Linear, Logistic, Survival, and Repeated Measures Models. 2nd ed. New York: Springer-Verlag, 2012:1272.
29. Wu Q, Zhong L, Li H, et al. A Follow-Up Study of Lung Function and Chest Computed Tomography at 6 Months after Discharge in Patients with Coronavirus Disease 2019. *Can Respir J* 2021;2021:6692409. doi: 10.1155/2021/6692409 [published Online First: 20210213]
30. Huang L, Yao Q, Gu X, et al. 1-year outcomes in hospital survivors with COVID-19: a longitudinal cohort study. *Lancet* 2021;398(10302):747-58. doi: 10.1016/S0140-6736(21)01755-4
31. Han X, Fan Y, Alwalid O, et al. Fibrotic Interstitial Lung Abnormalities at 1-year Follow-up CT after Severe COVID-19. *Radiology* 2021:210972. doi: 10.1148/radiol.2021210972 [published Online First: 20210727]
32. Han X, Fan Y, Alwalid O, et al. Six-month Follow-up Chest CT Findings after Severe COVID-19 Pneumonia. *Radiology* 2021;299(1):E177-E86. doi: 10.1148/radiol.2021203153 [published Online First: 20210126]
33. Stylemans D, Smet J, Hanon S, et al. Evolution of lung function and chest CT 6 months after COVID-19 pneumonia: Real-life data from a Belgian University Hospital. *Respir Med* 2021;182:106421. doi: 10.1016/j.rmed.2021.106421 [published Online First: 20210418]

34. Caruso D, Guido G, Zerunian M, et al. Post-Acute Sequelae of COVID-19 Pneumonia: Six-month Chest CT Follow-up. *Radiology* 2021;301(2):E396-E405. doi: 10.1148/radiol.2021210834 [published Online First: 20210727]
35. Guler SA, Ebner L, Aubry-Beigelman C, et al. Pulmonary function and radiological features 4 months after COVID-19: first results from the national prospective observational Swiss COVID-19 lung study. *Eur Respir J* 2021;57(4) doi: 10.1183/13993003.03690-2020 [published Online First: 2021/04/29]
36. Liu C, Ye L, Xia R, et al. Chest Computed Tomography and Clinical Follow-Up of Discharged Patients with COVID-19 in Wenzhou City, Zhejiang, China. *Ann Am Thorac Soc* 2020;17(10):1231-37. doi: 10.1513/AnnalsATS.202004-324OC
37. Tabatabaei SMH, Rajebi H, Moghaddas F, et al. Chest CT in COVID-19 pneumonia: what are the findings in mid-term follow-up? *Emerg Radiol* 2020;27(6):711-19. doi: 10.1007/s10140-020-01869-z [published Online First: 20201109]
38. Solomon JJ, Heyman B, Ko JP, et al. CT of Postacute Lung Complications of COVID-19. *Radiology* 2021:211396. doi: 10.1148/radiol.2021211396 [published Online First: 20210810]
39. Sonnweber T, Sahanic S, Pizzini A, et al. Cardiopulmonary recovery after COVID-19: an observational prospective multicentre trial. *Eur Respir J* 2021;57(4) doi: 10.1183/13993003.03481-2020 [published Online First: 2021/04/29]
40. Jacobi A, Chung M, Bernheim A, et al. Portable chest X-ray in coronavirus disease-19 (COVID-19): A pictorial review. *Clin Imaging* 2020;64:35-42. doi: 10.1016/j.clinimag.2020.04.001 [published Online First: 20200408]
41. DataSUS. MdS. SIGTAP - Sistema de Gerenciamento da Tabela de Procedimentos, Medicamentos e OPM do SUS. 2021 [cited 2021 January 03]. Available from: http://sigtap.datasus.gov.br/tabela-unificada/app/sec/inicio.jsp accessed January 03 2021.
42. Mettler FAM, Mahadevappa Bhargavan, MythreyiChambers, Charles E.Elee, Jennifer G. Frush, Donald P. Milano, Michael T. Miller, Donald L. Royal, Henry D. Spelic, David C.Ansari, Armin J. Bolch, Wesley E.Guebert, Gary M. Sherrier, Robert H. Smith, James M. Vetter, Richard J. Report Na. 184 - Medical Radiation Exposure of Patients in the United States. United States: National Council on Radiation Protection and Measurements, 2019.
43. Wilson J, Jungner G. Principles and Practice of Screening for Disease.: World Health Organization Public Health Papers, 1968.
44. FAPESP. COVID-19 DataSharing/BR 2021 [cited 2021 September 08]. Available from: https://repositoriodatasharingfapesp.uspdigital.usp.br/.

Figure Legends

Figure 1. Logistic regression-based machine learning model to detect the presence of COVID-19-related lung lesions. The patients were invited to participate in the study six months after COVID-19 positive RT-PCR at hospital admission. The modified Medical Research Council (mMRC) dyspnoea scale, oximetry (SpO2), spirometry (forced vital capacity [FVC]), and the five radiographic scores obtained during DL-based classification of CXR (p_{CXR}) were used as input data, and the presence of lung lesions due to COVID-19 was used as output data. AI: artificial intelligence. CT: computed tomography.

Figure 2. Flowchart of patient selection. FVC, forced vital capacity; LLN, lower limit of normal; mMRC, modified Medical Research Council dyspnoea scale. *Rest SpO2 < 90% or a decrease in SpO2 of at least 4% after the 1-min sit and stand test.

Figure 3. Fibrotic-like changes after critical COVID-19 in a patient in his early 70s. (A) PA chest radiograph obtained 7 months after infection shows reticular opacities with a slight peripheral predominance diffusely distributed in both lungs. (B) Image from the same radiograph analysed by the AI algorithm with a heat map highlighting the areas of pulmonary involvement. (C, D) Chest CT obtained 8 months after infection shows moderate ground glass opacities, linear multifocal and reticular abnormalities, discrete traction bronchiectasis and slight parenchymal architectural distortion. The patient had dyspnoea (mMRC=1) and altered FVC (2.34 L / 60% pred), besides the normal oximetry (97%).

Figure 4. Flowchart for lung lesion case-finding in COVID-19 survivors. *Altered oximetry: Resting SpO2 ≤90% or a decrease in SpO2 of ≥4% during the 1-min sit and stand test. **Altered CXR: COVID-19 findings, including bilateral linear and/or reticular opacities, especially peripheral opacities. † The in-person consultation also should start with oximetry and mMRC examinations. †† The suggestion is to perform plethysmography with diffusion capacity measure. CXR, chest X-Ray; FVC, forced vital capacity; LLN, lower limit of normal; mMRC, modified Medical Research Council dyspnoea scale.

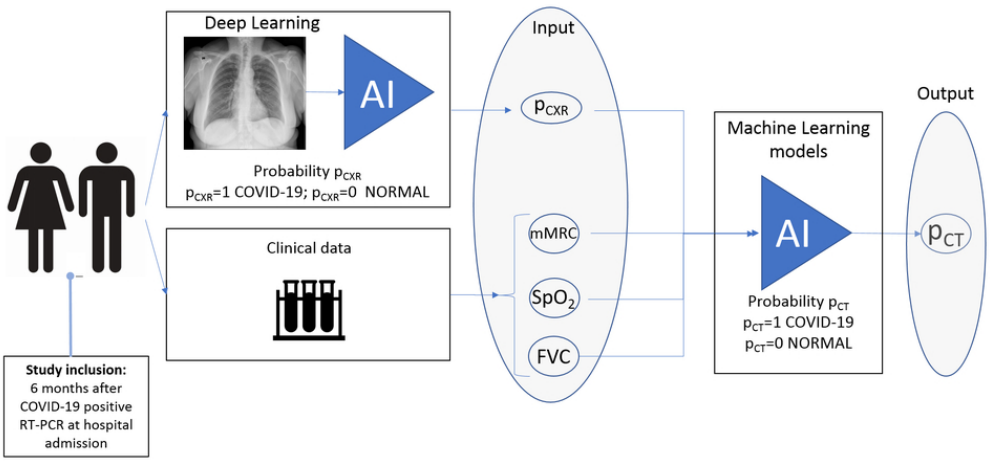


Figure 1. Logistic regression-based machine learning model to detect the presence of COVID-19-related lung lesions. The patients were invited to participate in the study six months after COVID-19 positive RT-PCR at hospital admission. The modified Medical Research Council (mMRC) dyspnoea scale, oximetry (SpO2), spirometry (forced vital capacity [FVC]), and the five radiographic scores obtained during DL-based classification of CXR (pCXR) were used as input data, and the presence of lung lesions due to COVID-19 was used as output data. AI: artificial intelligence. CT: computed tomography.

80x37mm (300 x 300 DPI)

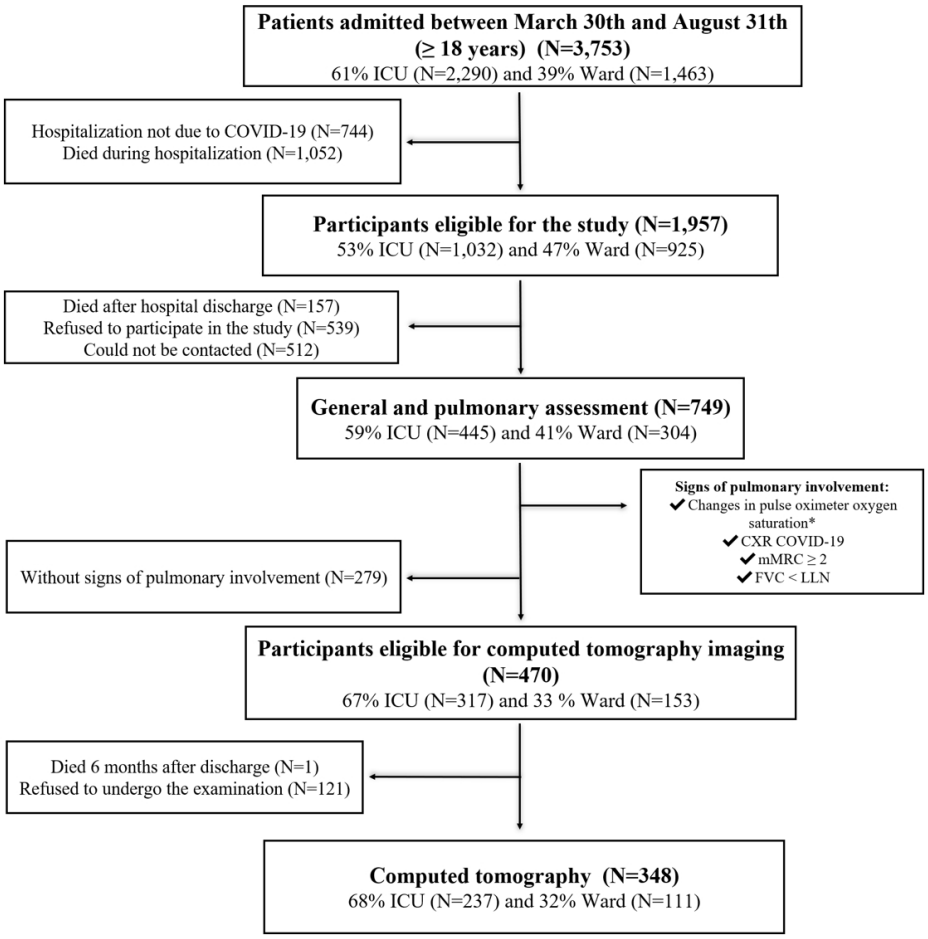


Figure 2. Flowchart of patient selection. FVC, forced vital capacity; LLN, lower limit of normal; mMRC, modified Medical Research Council dyspnoea scale. *Rest SpO2 < 90% or a decrease in SpO2 of at least 4% after the 1-min sit and stand test.

51x47mm (600 x 600 DPI)

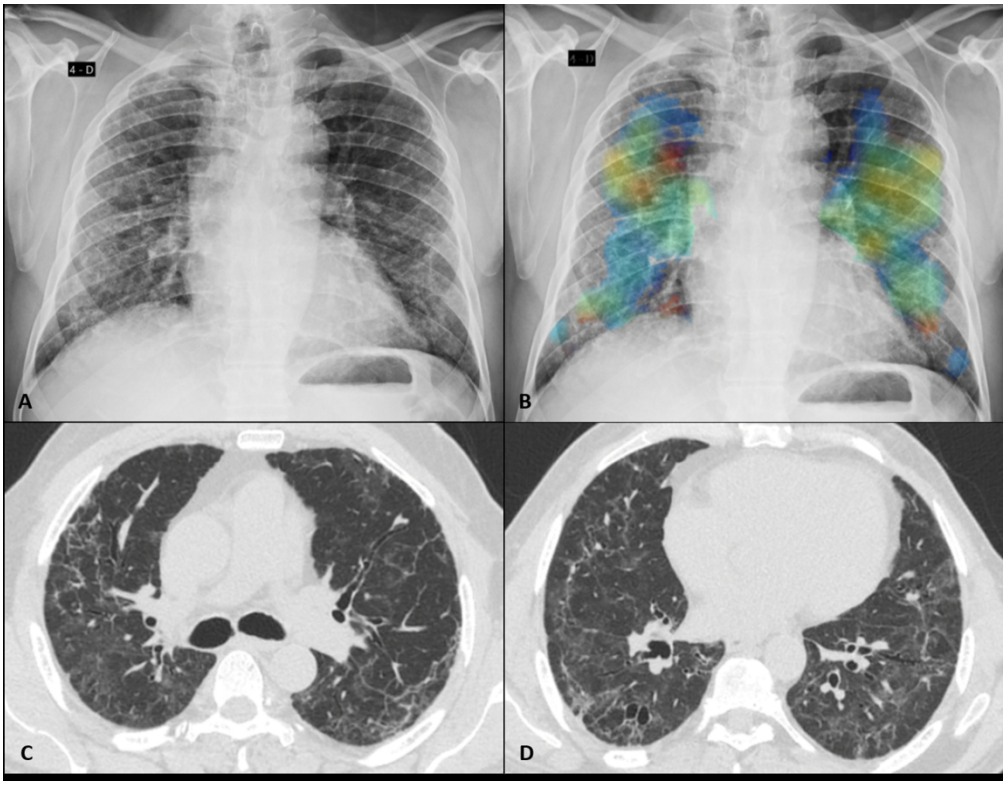


Figure 3. Fibrotic-like changes after critical COVID-19 in a patient in his early 70s. (A) PA chest radiograph obtained 7 months after infection shows reticular opacities with a slight peripheral predominance diffusely distributed in both lungs. (B) Image from the same radiograph analysed by the AI algorithm with a heat map highlighting the areas of pulmonary involvement. (C, D) Chest CT obtained 8 months after infection shows moderate ground glass opacities, linear multifocal and reticular abnormalities, discrete traction bronchiectasis and slight parenchymal architectural distortion. The patient had dyspnoea (mMRC=1) and altered FVC (2.34 L / 60% pred), besides the normal oximetry (97%).

154x119mm (600 x 600 DPI)

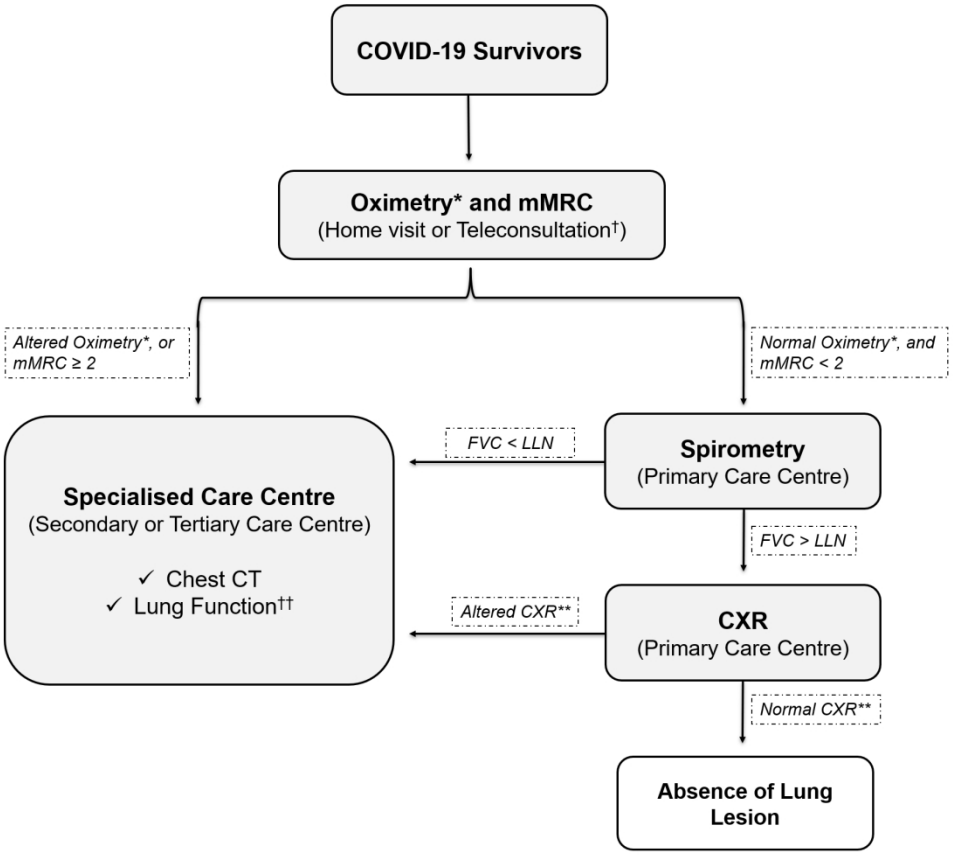


Figure 4. Flowchart for lung lesion case-finding in COVID-19 survivors. *Altered oximetry: Resting SpO2 ≤90% or a decrease in SpO2 of ≥4% during the 1-min sit and stand test. **Altered CXR: COVID-19 findings, including bilateral linear and/or reticular opacities, especially peripheral opacities. † The in-person consultation also should start with oximetry and mMRC examinations. †† The suggestion is to perform plethysmography with diffusion capacity measure. CXR, chest X-Ray; FVC, forced vital capacity; LLN, lower limit of normal; mMRC, modified Medical Research Council dyspnea scale.

81x71mm (600 x 600 DPI)

Supplemental Material

Chronic lung lesions in COVID-19 survivors: predictive clinical model

Carlos R R Carvalho, Rodrigo C Chate, Marcio VY Sawamura, Michelle L Garcia, Celina A Lamas, Diego AC Cardenas, Daniel M Lima, Paula G Scudeller, João M Salge, Cesar H Nomura, Marco A Gutierrez, HCFMUSP Covid-19 Study Group.

Contents

Supplemental Methods

Datasets

The SIIM-RSNA dataset contains 6,334 posterior-anterior radiographic images from 6,054 patients obtained from the public dataset Machine Learning Challenge on COVID-19 Pneumonia Detection and Localization.[1] Specialists classified images as “negative for pneumonia” or “COVID-19 pneumonia”. A total of 6,030 images were selected and randomly distributed in training and validation sets (1,276 negative and 3,711 positive findings) and a test set (400 negative and 643 positive findings).

The Institute of Radiology (InRad) dataset contains chest X-Ray (CXR) and chest computed tomographic (CT) images of 257 patients. The CXR images were classified as normal (n=145) or with findings related to COVID-19 (n=112) and randomly distributed in training and validation sets (214 patients) and a test set (n=43). Images were obtained from the InRad of the Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo (HCFMUSP).

Because of differences in dataset sizes, a data augmentation technique was adopted using random transformations, including rotation (0–15 degrees), horizontal mirroring, and random changes in intensity and contrast (0–5%).

Classification of chest radiography images

A deep-learning (DL) approach using a convolutional neural network (CNN) based on an EfficientNetB7 architecture was used.[2] The network classification layer was replaced by a global average pooling operation, followed by batch normalization and the adoption of a dense layer with one neuron and sigmoid activation function. Each training iteration was run for 40 epochs with an Adam optimizer at a learning rate of 0.0001. All images were resized to 600 x 600 pixels.

The CNN was trained using the SIIM-RSNA dataset to detect radiographic patterns of COVID-19 pneumonia. Training was initiated in EfficientNetB7 using weights after pre-training with the ImageNet dataset.[3]

A five-fold cross-validation strategy was adopted for the training and validation sets. The training weights obtained for each fold were used with the

test set of the SIIM-SNA to evaluate classification accuracy (Table 1). The fold with the best area under the receiver operating characteristic curve (AUC), in this case, fold 1 with AUC of 0.89, defines the final weights of the CNN.

Table 1. Classification of the test set of the SIIM-RSNA dataset as negative (normal) or positive (patterns of COVID-19 pneumonia).

Dataset	5-fold	Acc	Prec	Sensitivity	Specificity	F1-score	AUC
SIIM-RSNA	0	0.80	0.85	0.82	0.76	0.83	0.88
	1	0.80	0.85	0.82	0.77	0.84	**0.89**
	2	0.78	0.77	0.92	0.56	0.84	0.87
	3	0.76	0.74	0.93	0.48	0.83	0.86
	4	0.76	0.74	0.93	0.48	0.83	0.86

Area under the receiver operating characteristic curve (AUC); Accuracy (Acc); Precision (Prec).

For the InRad dataset, the CNN was initialized with the final weights defined in the training set of SIIM-RSNA. After initialization, the CNN was retrained to classify images as normal or with findings related to COVID-19.

The InRad dataset was divided into six-folds during the retraining, five folds for training and validation, and one-fold for test. To avoid bias, the test fold was selected to run all six folds available and, for each test fold selected, a five-fold cross-validation strategy was applied in the remaining training and validation folds (Table 2).

Table 2. Classification using six test folds of the InRad database.

Dataset	Test fold	Acc	Prec	Sensitivity	Specificity	F1-score	AUC
InRad	0	0.79±0.01	0.74±0.04	0.82±0.07	0.77±0.06	0.78±0.02	0.86±0.02
	1	0.69±0.02	0.62±0.03	0.84±0.06	0.57±0.07	0.71±0.02	0.75±0.01
	2	0.67±0.05	0.60±0.06	0.81±0.08	0.57±0.13	0.68±0.02	0.76±0.02
	3	0.77±0.04	0.71±0.07	0.80±0.04	0.74±0.10	0.75±0.03	0.80±0.02
	4	0.82±0.05	0.77±0.11	0.89±0.10	0.78±0.14	0.81±0.03	0.89±0.04
	5	0.71±0.04	0.62±0.04	0.90±0.02	0.58±0.08	0.73±0.03	0.80±0.02

Data represent the mean and standard deviation after five-fold cross validation. Area under the receiver operating characteristic curve (AUC); Accuracy (Acc); Precision (Prec).

Detection of chronic lung lesions on computed tomography images

Three machine learning models were developed based on the clinical data, including the modified Medical Research Council dyspnea scale (mMRC), oximetry (SpO_2) and spirometry (forced vital capacity, FVC), and five radiographic

probabilities (p_{CXR0} top_{CXR4}) with findings related to COVID-19 (p_{CXRn}=1) and normal (p_{CXRn}=0), which were obtained from the previous step (Table 2). As output, the models predict the value of a binary variable (p_{CT}) related to the presence of chronic lung lesions on CT images, with p_{CT}=1 for a CT score ≥ 7 (n=129) and p_{CT}=0 for a CT score < 7 (n=128) (Figure 1).

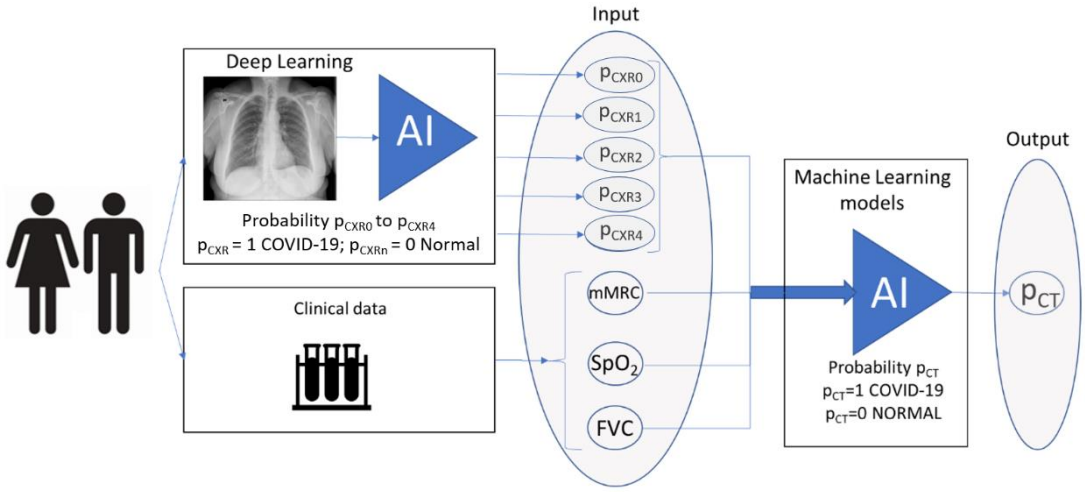


Figure 1. Machine learning-based model. Data on the modified Medical Research Council (mMRC) dyspnea scale, oximetry (SpO2), and spirometry (forced vital capacity [FVC]), and radiographic probabilities (p_{CXR0} to p_{CXR4}) with findings related to COVID-19 (p_{CXRn}=1) and normal (p_{CXRn}=0) were used as input variables, and the presence of lung lesions due to COVID-19 (p_{CT}) was used as output. AI, artificial intelligence. CT, computed tomography.

The first model was a logistic regression (LR) model with L2 regularization to prevent overfitting,[4] whereas the second model was a random forest model with 100 trees (RF-100), Gini criterion, minimum of two samples for splitting, minimum of one sample in leaves, and bootstrap.[4] The third model was a random forest model with parameters as described above, except for the limit of 10 trees and maximum depth h_max=6 (RF-10).[4] The performance of the machine-learning models was evaluated based on sensitivity, specificity, AUC, and F1-score.

Three combinations of input variables were evaluated: 1) clinical variables (mMRC, SpO_2, and FVC); 2) CXR; and 3) clinical variables (mMRC, SpO_2, FVC) and CXR.

The performance of the LR model was better when a combination of all variables (clinical variables and CXR) was used. The following metrics expressed in terms of mean ± standard deviation and 95% Confidence Interval (CI) were considered: sensitivity, 0.85±0.08 (95% CI [0.77, 0.94]); specificity, 0.70±0.14 (95% CI [0.55, 0.85]); F1-score, 0.79±0.06 (95% CI [0.73, 0.85]); and AUC, 0.80±0.07(95% CI [0.72, 0.87]) (Table 3).

Table 3. Predictive performance of three multivariate models using three datasets.

Groups of variables	Method	Sensitivity	Specificity	F1-score	AUC
1 SpO_2, mMRC score, and FVC	LR	0.87±0.16	0.42±0.33	0.71±0.03	0.68±0.10
	RF-10	0.88±0.15	0.37±0.32	0.71±0.03	0.66±0.08
	RF-100	0.82±0.12	0.44±0.13	0.69±0.08	0.62±0.12
2 CXR	LR	0.88±0.05	0.52±0.14	0.75±0.04	0.78±0.05
	RF-10	0.91±0.08	0.41±0.18	0.73±0.04	0.73±0.06
	RF-100	0.94±0.07	0.33±0.19	0.72±0.03	0.72±0.03
3 SpO_2, mMRC score, FVC and CRX	LR	**0.85±0.08**	**0.70±0.14**	**0.79±0.06**	**0.80±0.07**
	RF-10	0.85±0.09	0.61±0.22	0.76±0.04	0.76±0.08
	RF-100	0.89±0.06	0.49±0.17	0.75±0.04	0.76±0.07

Values are presented as the mean ± standard deviation after five-fold cross validation for each test fold. Area under the receiver operating characteristic curve (AUC); Accuracy (Acc); Chest X-Ray (CRX); Forced vital capacity (FVC); Logistic Regression (LR); modified Medical Research Council dyspnea scale (mMRC); Precision (Prec); Random forest (RF).

The LR model is represented by the following function:

$$p_{CT} = \sigma\left(\beta_1 FVC^* + \beta_2\, mMRC^* + \beta_3 S_p O_2 + \beta_4 p_{CXR0} + \beta_5 p_{CXR1} + \beta_6 p_{CXR2} + \beta_7 p_{CXR3} + \beta_8 p_{CXR4}\right)$$

$$\beta_1 = -0.3705 \;\; \beta_2 = -2.2807 \;\; \beta_3 = -0.745 \;\; \beta_4 = 1.1257$$

$$\beta_5 = 1.4960 \;\; \beta_6 = 1.0761 \;\; \beta_7 = 0.7328 \;\; \beta_8 = -0.7613$$

where p_{CT} is the probability of the presence of abnormalities on CT images, σ is the sigmoid function to restrict p_{CT} between 0 and 1, $FVC^* = \frac{FVC_{Resting}}{2FVC_{min}}$, $mMRC^* = \frac{mMRC}{4}$, and p_{CXR0} to p_{CXR4} are the probabilities that the CXR image has findings related to sequelae from COVID-19, obtained in each fold (0 to 4) during a 5-folds cross validation. Table 4 shows the estimates for the logistic regression function.

Table 4. Estimates of the logistic regression function.						
Variable	**Estimated regression coefficient (β)**	**Estimated Standard Error**	***p*-value**	**95% CI for regression coefficient (β)**		**Estimated odds ratios**
FVC^*	-0.3705	0.3210	0.248	-0.9990	0.2580	0.6904
$mMRC^*$	-2.2807	0.3020	<0.001	-2.8730	-1.6890	0.1022
S_pO_2	-0.7450	0.2320	0.001	-1.2010	-0.2890	0.4747
p_{CXR0}	1.1257	0.4150	0.007	0.3120	1.9400	3.0824
p_{CXR1}	1.4960	0.4160	<0.001	0.6810	2.3110	4.4638
p_{CXR2}	1.0761	0.3390	0.002	0.4120	1.7410	2.9332
p_{CXR3}	0.7328	0.3380	0.030	0.0710	1.3950	2.0809
p_{CXR4}	-0.7613	0.4580	0.096	-1.6590	0.1360	0.4671
Forced vital capacity (FVC); modified Medical Research Council dyspnea scale (mMRC); radiographic probabilities (Pcxr0 to Pcxr4).						

Also, we included demographic and anthropometric variables on the logistic regression prediction model, performing experiments using six different combinations of variables (age, gender, body mass index [BMI], SpO2, mMRC score, FVC and CXR). The performance of each combination is reported in the Table 5. The model performance with the inclusion of demographic or anthropometric variables did not result in significant improvement. According to our experiments, the combination of SpO2, mMRC score, FVC and CXR presented the best performance.

Table 5. Performance of the predictive model using six combinations of variables (N=257).				
Groups of variables	**Sensitivity**	**Specificity**	**F1-score**	**AUC**
1 Age, gender, and BMI	0.87±0.09	0.40±0.27	0.71±0.03	0.64±0.09
2 SpO_2, mMRC score, and FVC	0.87±0.16	0.42±0.33	0.71±0.03	0.68±0.10
3 Age, Gender, BMI, SpO2, mMRC score, and FVC	0.95±0.05	0.37±0.30	0.75±0.06	0.71±0.10
4 CXR	0.88±0.05	0.52±0.14	0.75±0.04	0.78±0.05
5 Age, Gender, BMI, SpO_2, mMRC score, FVC, and CXR	0.87±0.08	0.65±0.16	0.79±0.06	0.79±0.06
6 SpO_2, mMRC score, FVC, and CXR	0.85±0.08	0.70±0.14	0.79±0.06	0.80±0.07
Values are presented as the mean ± standard deviation after five-fold cross validation for each test fold. Area under the receiver operating characteristic curve (AUC); Body Mass Index (BMI); Chest X-Ray (CRX); Forced vital capacity (FVC); modified Medical Research Council dyspnoea scale (mMRC).				

Dataset and normalization of clinical data

A total of 257 patients with data on the mMRC dyspnea scale, oximetry, spirometry, CRX, and chest CT were selected to predict pulmonary changes. Of the 257 patients, 128 had no significant CT changes (scores < 7). A CT score of 7 was used as the cutoff value by maximizing F1 scores and AUC (Figure 2).

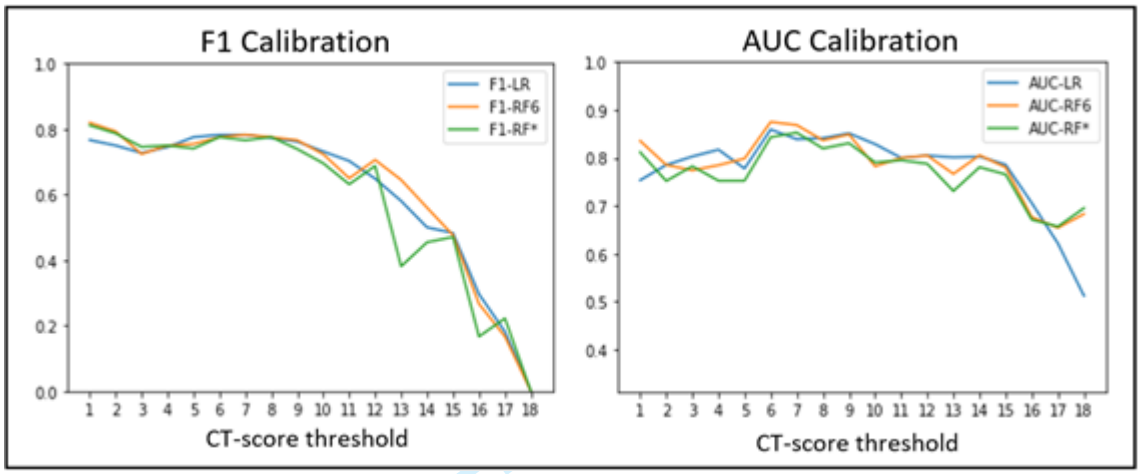
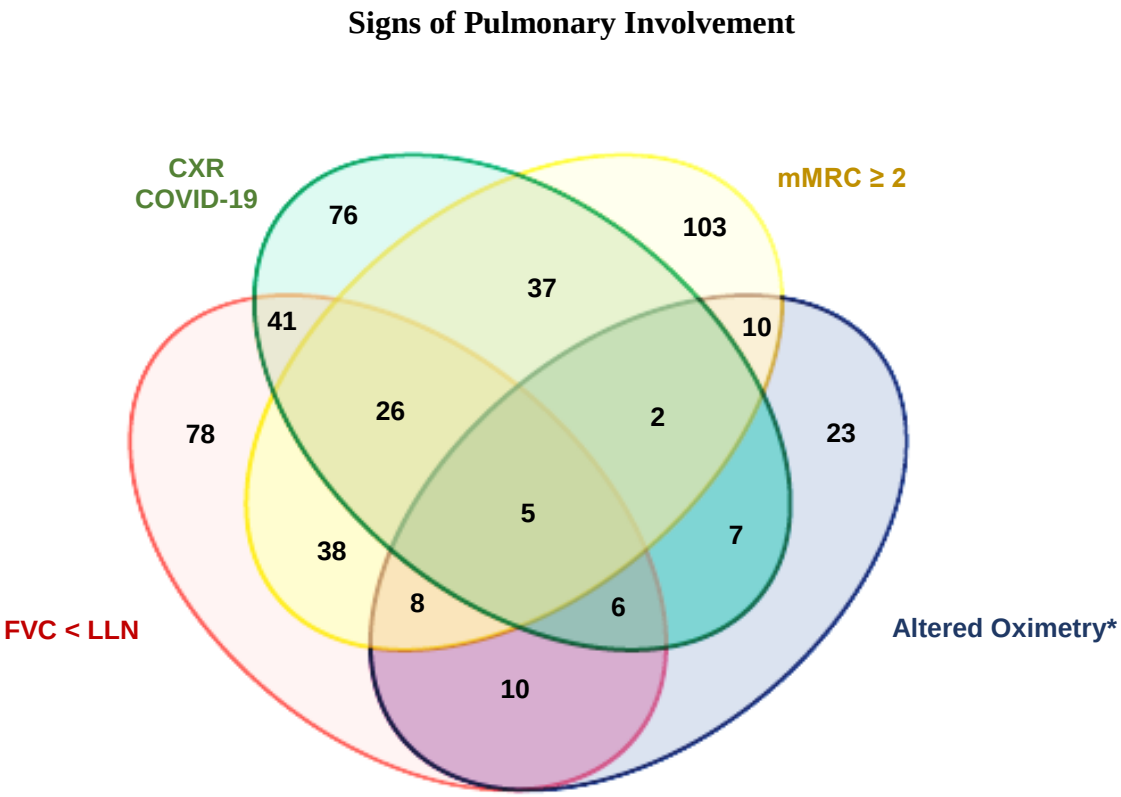
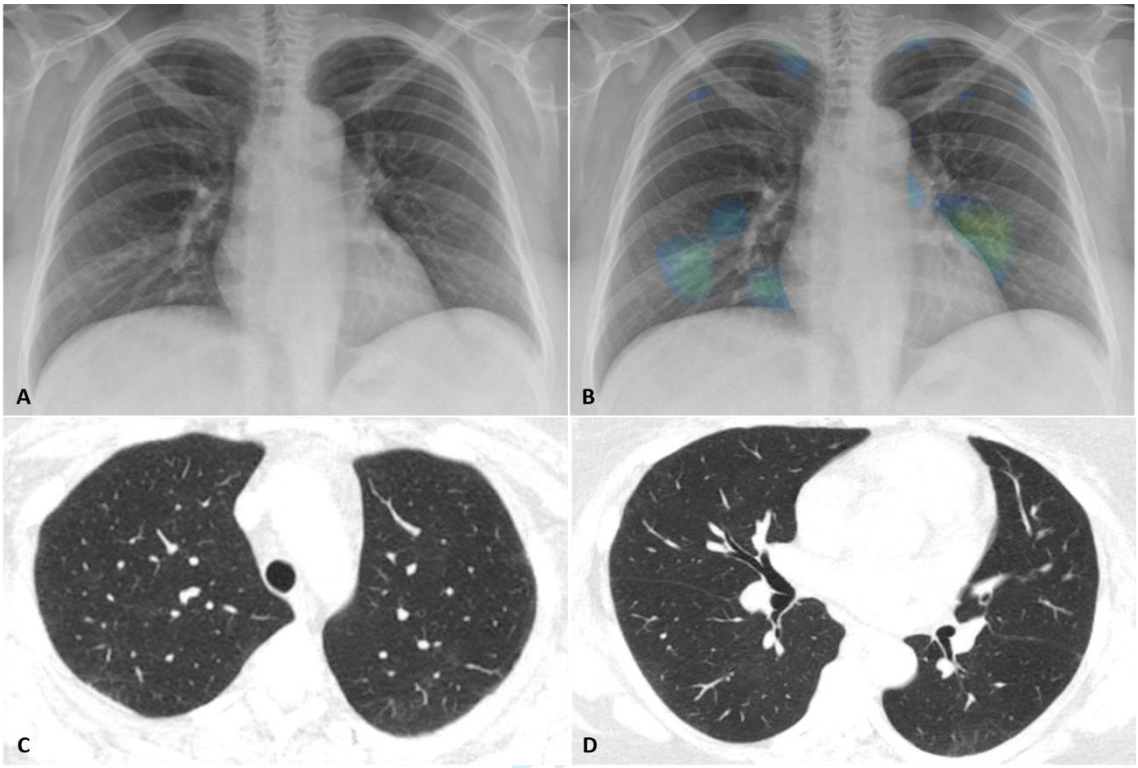


Figure 2. Computed tomography scores based on the F1-score and AUC values.

Clinical variables were normalized by dividing the mMRC values by 4 (resulting in values between 0 and 1) and the $FVC_{Resting}$ by twice the FVC_{min} (resulting in a minimum value of 0.257 and a maximum value of 0.847).



Supplemental Figure S1. Diagram showing the overlap in the changes of parameters used as pulmonary criteria to refer patients for thorax computed tomography. Values are expressed as the number of patients showing the correspondent alterations. CXR, chest X-Ray; FVC, forced vital capacity; LLN, lower limit of normal; mMRC, modified Medical Research Council dyspnea scale. *Resting SpO_2 ≤ 90% or a decrease in SpO_2 of ≥ 4% during the 1 min sit-and-stand test.



Supplemental Figure 2. Representative scan of a patient in her late 40s showing resolving ground glass abnormality after moderate COVID-19. (A) PA chest radiograph obtained 8 months after admission was considered normal by radiologists. (B) The same radiograph analyzed by the AI algorithm with heat map. Small focal abnormalities in the apical and paracardiac regions of the lungs are highlighted in green and blue. (C, D) Chest CT obtained 11 months after admission shows mild residual ground glass abnormality in the periphery of the upper lobes and left lower lobe. The patient complained of dyspnea (mMRC=3) but had normal lung function (FVC=3.81 L/91% pred) and normal oximetry (99%).

Supplemental Table S1. Demographic and clinical characteristics of the cohort of post-COVID-19 patients in this study (N=749).

Variables	Values
Age (years)	56.1 (44.4–65.1)
Male sex	399 (53.3)
BMI (kg/m^2)	30.8 (27.7–35.6) {746}
Comorbidities	
Hypertension	425 (56.7)
Smokers	285/743 (38.4)
Diabetes	261 (34.8)
COPD	55 (7.3)
Admission	
ICU	445 (59.4)
Length of ICU stay (days)	10 (6–18) {445}
IMV	304/445 (68.3)
Vital signs	
Body temperature (°C)	36.1 (35.6–36.0) {748}
Systolic blood pressure (mmHg)	124 (116–135) {743}
Diastolic blood pressure (mmHg)	77 (70–84) {743}
Heart rate (bpm)	73 (67–83) {747}
Respiratory rate (rpm)	20 (18–2) {736}
Oxygen saturation (%)	97 (95.2–98) {746}

Values are presented as median (IQR), median (IQR) {n}, n (%), or n/N (%). COPD, chronic obstructive pulmonary disease; BMI, body mass index; ICU, intensive care unit. IMV, invasive mechanical ventilation.

Supplemental Table S2. Demographic and clinical characteristics of patients with and without pulmonary involvement (N=749).

Variables	Pulmonary involvement (n=470)	No pulmonary involvement (n=279)	*p*-value
Age (years)	57.9 (45.7–65.8)	53.9 (42.5–63.7)	0.000
Male sex	228 (48.5)	171 (61.3)	0.001
BMI (kg/m^2)	31.2 (27.7–35.9) {469}	30.5 (27.6–35.2) {277}	0.111
Comorbidities			
Hypertension	287 (61.1)	138 (49.5)	0.000
Smokers	188/468 (40.2)	97/275 (35.3)	0.104
Diabetes	179 (38.1)	82 (29.4)	0.009
COPD	42 (8.9)	13 (4.7)	0.044
Admission			
ICU	317 (67.4)	128 (45.9)	0.000
Length of ICU stay (days)	11 (6–20) {317}	8 (4–14) {128}	0.000
IMV	222/317 (70)	82/128 (64.1)	0.260

Values are presented as median (IQR), median (IQR) {n}, n (%), or n/N (%). COPD, chronic obstructive pulmonary disease; BMI, body mass index; ICU, intensive care unit. IMV, invasive mechanical ventilation.

Supplemental Table S3. Demographic and clinical characteristics of COVID-19 patients with signs of pulmonary involvement (N=470).

Variables	Patients with signs of pulmonary involvement		*p*-value
	Those who underwent CT (n=348)	Those who did not undergo CT (n=122)	
Age (years)	57.8 (45.7–65.8)	58.1 (45.3–65.8)	0.490
Male sex	163 (46.8)	65 (53.3)	0.392
BMI (kg/m^2)	31.6 (28.0–36.0)	30.3 (27.0–35.9) {121}	0.041
Comorbidities			
Hypertension	215 (61.8)	72 (59)	0.469
Smokers	139/347 (40.1)	49/121 (40.5)	0.762
Diabetes	142 (40.8)	37 (30.3)	0.999
COPD	32 (9.2)	10 (8.2)	0.826
Admission			
ICU	237 (68.1)	80 (65.6)	0.999
Length of ICU stay (days)	11 (6–20) {237}	10 (4.7–19) {80}	0.913
IMV	174/237 (73.4%)	48/80 (60%)	0.034

Values are presented as median (IQR), median (IQR) {n}, n (%), or n/N (%). COPD, chronic obstructive pulmonary disease; BMI, body mass index; ICU, intensive care unit. IMV, invasive mechanical ventilation.

Supplemental Table S4. Chest computed tomography (CT) features in COVID-19 patients with CT score ≥ 7 (N=156).	
Variables	**CT changes**
CT score ≥ 7	156/328 (47.6)
Characteristics (n=156)	
Ground-glass opacities	153 (98.1)
Parenchymal bands	143 (91.7)
Reticulations	134 (85.9)
Traction bronchiectasis	92 (59)
Architectural distortion	73 (46.8)
Perilobular opacities	50 (32.1)
Bronchial wall thickening	38 (24.4)
Mosaic attenuation pattern	32 (20.5)
Consolidations	3 (1.9)
Pneumatocele	2 (1.3)
Honeycombing	-
Of the 328 patients who underwent CT scan, 47.6% had a CT score ≥ 7. Values are n/N (%) or n (%).	

Supplemental Table S5. Computed tomography changes 6 to 11 months after hospitalization due to COVID-19 (N=328).

Characteristics	Total cohort (N=328)	ICU Patients (N=222)	Ward Patients (N=106)
Ground-glass opacities	251 (76.5)	197 (86.6)	54 (51.3)
Parenchymal bands	209 (63.7)	169 (76.5)	40 (41)
Reticulations	169 (51.5)	145 (66.5)	24 (23.1)
Traction bronchiectasis	98 (29.9)	91 (44.1)	7 (7.7)
Architectural distortion	78 (23.8)	73 (35.8)	5 (6.4)
Bronchial wall thickening	89 (27.1)	60 (27.4)	29 (25.6)
Mosaic attenuation pattern	58 (17.7)	46 (20.1)	12 (11.5)
Perilobular opacities	50 (14)	47 (24.6)	3 (2.6)
Consolidation	3 (0.9)	3 (1.7)	-
Pneumatocele	2 (0.6)	2 (1.1)	-
Honeycombing	-	-	-

Values are presented as n (%).

Supplemental Table S6. Demographic and clinical characteristics of COVID-19 patients with pulmonary involvement stratified by inclusion in prediction analysis of pulmonary changes (N=328).

Variables	Patients with Pulmonary Changes		*p*-value
	Included Patients (N=257)	Excluded Patients (N=91)	
Age (years)	56.5 (45.7–64.4)	60.5 (46.9–69.9)	0.011
Male sex	113 (44)	50 (54.9)	0.068
BMI (kg/m^2)	32 (28.8–36.8)	30.6 (26.8–35.4)	0.054
Comorbidities			
Hypertension	151 (58.7)	64 (70.3)	0.060
Smokers	97/256 (37.9)	42 (46.1)	0.173
Diabetes	103 (40.1)	39 (42.9)	0.710
COPD	20 (7.8)	12 (13.2)	0.141
Admission			
ICU	179 (69.6)	58 (63.7)	0.359
Length of ICU stay (days)	12 (6–20.5) {179}	9.5 (6.2–19.7) {58}	0.209
IMV	140 (54.7)	35 (38.6)	0.010

Values are presented as median (IQR), median (IQR) {n}, n (%), or n/N (%). BMI, body mass index; COPD, chronic obstructive pulmonary disease; ICU, intensive care unit. IMV, invasive mechanical ventilation.

Supplemental References

1. Stephens K. SIIM, FISABIO, and RSNA Host Machine Learning Challenge for COVID-19 Detection and Localization. . *AXIS Imaging News* . 2021.

2. Tan M, Le Q. Efficientnet: Rethinking model scaling for convolutional neural networks. *International Conference on Machine Learning* 2019. p. 6105-14.

3. Russakovsky O, Deng J, Su H, et al. ImageNet Large Scale Visual Recognition Challenge. *International Journal of Computer Vision* 2015; **115**: 211-52.

4. Vittinghoff E, Glidden DV, Shiboski SC, et al. Regression Methods in Biostatistics: Linear, Logistic, Survival, and Repeated Measures Models. 2nd ed. New York: Springer-Verlag, 2012:1272.

TRIPOD Checklist: Prediction Model Development and Validation

Section/Topic	Item		Checklist Item	Page
Title and abstract				
Title	1	D;V	Identify the study as developing and/or validating a multivariable prediction model, the target population, and the outcome to be predicted.	1
Abstract	2	D;V	Provide a summary of objectives, study design, setting, participants, sample size, predictors, outcome, statistical analysis, results, and conclusions.	2
Introduction				
Background and objectives	3a	D;V	Explain the medical context (including whether diagnostic or prognostic) and rationale for developing or validating the multivariable prediction model, including references to existing models.	4
	3b	D;V	Specify the objectives, including whether the study describes the development or validation of the model or both.	5
Methods				
Source of data	4a	D;V	Describe the study design or source of data (e.g., randomized trial, cohort, or registry data), separately for the development and validation data sets, if applicable.	5
	4b	D;V	Specify the key study dates, including start of accrual; end of accrual; and, if applicable, end of follow-up.	5
Participants	5a	D;V	Specify key elements of the study setting (e.g., primary care, secondary care, general population) including number and location of centres.	5
	5b	D;V	Describe eligibility criteria for participants.	5
	5c	D;V	Give details of treatments received, if relevant.	-
Outcome	6a	D;V	Clearly define the outcome that is predicted by the prediction model, including how and when assessed.	7
	6b	D;V	Report any actions to blind assessment of the outcome to be predicted.	Data Supplement (Pg. 3 and 5)
Predictors	7a	D;V	Clearly define all predictors used in developing or validating the multivariable prediction model, including how and when they were measured.	6, 7
	7b	D;V	Report any actions to blind assessment of predictors for the outcome and other predictors.	Data Supplement (Pg. 3 and 5)
Sample size	8	D;V	Explain how the study size was arrived at.	5
Missing data	9	D;V	Describe how missing data were handled (e.g., complete-case analysis, single imputation, multiple imputation) with details of any imputation method.	Data Supplement, (Pg. 6)
Statistical analysis methods	10a	D	Describe how predictors were handled in the analyses.	Data Supplement (Pg. 3 and 5)
	10b	D	Specify type of model, all model-building procedures (including any predictor selection), and method for internal validation.	Data Supplement (Pg. 3 and 5)
	10c	V	For validation, describe how the predictions were calculated.	Data Supplement (Pg. 3 and 5)
	10d	D;V	Specify all measures used to assess model performance and, if relevant, to compare multiple models.	Data Supplement (Pg. 3 and 5)
	10e	V	Describe any model updating (e.g., recalibration) arising from the validation, if done.	S Data Supplement (Pg. 4)
Risk groups	11	D;V	Provide details on how risk groups were created, if done.	n.a.
Development vs. validation	12	V	For validation, identify any differences from the development data in setting, eligibility criteria, outcome, and predictors.	Data Supplement (Pg. 3 and 5)
Results				
Participants	13a	D;V	Describe the flow of participants through the study, including the number of participants with and without the outcome and, if applicable, a summary of the follow-up time. A diagram may be helpful.	8
	13b	D;V	Describe the characteristics of the participants (basic demographics, clinical features, available predictors), including the number of participants with missing data for predictors and outcome.	8
	13c	V	For validation, show a comparison with the development data of the distribution of important variables (demographics, predictors and outcome).	Data Supplement (Pg. 3 and 5)
Model development	14a	D	Specify the number of participants and outcome events in each analysis.	Data Supplement (Pg. 3 and 5)
	14b	D	If done, report the unadjusted association between each candidate predictor and outcome.	Data Supplement (Pg. 3 and 5)
Model specification	15a	D	Present the full prediction model to allow predictions for individuals (i.e., all regression coefficients, and model intercept or baseline survival at a given time point).	Data Supplement (Pg. 3 and 5)

TRIPOD Checklist: Prediction Model Development and Validation

	15b	D	Explain how to the use the prediction model.	Data Supplement (Pg. 3 and 5)
Model performance	16	D;V	Report performance measures (with CIs) for the prediction model.	Data Supplement (Pg. 5)
Model-updating	17	V	If done, report the results from any model updating (i.e., model specification, model performance).	Data Supplement (Pg.5)
Discussion				
Limitations	18	D;V	Discuss any limitations of the study (such as nonrepresentative sample, few events per predictor, missing data).	12
Interpretation	19a	V	For validation, discuss the results with reference to performance in the development data, and any other validation data.	Data Supplement (Pg.5)
	19b	D;V	Give an overall interpretation of the results, considering objectives, limitations, results from similar studies, and other relevant evidence.	10, 11, 12
Implications	20	D;V	Discuss the potential clinical use of the model and implications for future research.	10, 11, 12
Other information				
Supplementary information	21	D;V	Provide information about the availability of supplementary resources, such as study protocol, Web calculator, and data sets.	n.a.
Funding	22	D;V	Give the source of funding and the role of the funders for the present study.	14

*Items relevant only to the development of a prediction model are denoted by D, items relating solely to a validation of a prediction model are denoted by V, and items relating to both are denoted D;V. We recommend using the TRIPOD Checklist in conjunction with the TRIPOD Explanation and Elaboration document.

BMJ Open

Chronic lung lesions in COVID-19 survivors: predictive clinical model

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2021-059110.R3
Article Type:	Original research
Date Submitted by the Author:	07-Mar-2022
Complete List of Authors:	Carvalho, Carlos; Universidade de São Paulo, Instituto do Coração - Divisão de Pneumologia Chate, Rodrigo; Universidade de São Paulo Hospital das Clínicas, Instituto de Radiologia Sawamura, Marcio; Universidade de Sao Paulo Hospital das Clinicas, Instituto de Radiologia Garcia, Michelle; Universidade de São Paulo Hospital das Clínicas, Instituto do Coração - Divisão de Pneumologia Lamas, Celina ; Universidade de Sao Paulo Hospital das Clinicas, Instituto do Coração - Divisão de Pneumologia Cardenas, Diego; Universidade de São Paulo Hospital das Clínicas, Instituto do Coração - Divisão de Informática Lima, Daniel Mario; Universidade de São Paulo Hospital das Clínicas, Instituto do Coração - Divisão de Informática Scudeller, Paula; Universidade de São Paulo Hospital das Clínicas, Instituto do Coração - Divisão de Pneumologia Salge, João; Universidade de São Paulo Hospital das Clínicas, Instituto do Coração - Divisão de Pneumologia Nomura, Cesar; Universidade de São Paulo Hospital das Clínicas, Instituto de Radiologia Gutierrez, Marco; Universidade de São Paulo Hospital das Clínicas, Instituto do Coração - Divisão de Pneumologia
Primary Subject Heading:	Respiratory medicine
Secondary Subject Heading:	Global health
Keywords:	COVID-19, Chest imaging < RADIOLOGY & IMAGING, RESPIRATORY MEDICINE (see Thoracic Medicine)

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Chronic lung lesions in COVID-19 survivors: predictive clinical model

Carlos R R Carvalho (0000-0002-1618-8509)[1], Rodrigo C Chate[2], Marcio VY Sawamura[2], Michelle L Garcia[1], Celina A Lamas[1], Diego AC Cardenas[3], Daniel M Lima[3], Paula G Scudeller[1], João M Salge[1], Cesar H Nomura[2], Marco A Gutierrez[3], HCFMUSP Covid-19 Study Group*

1 Pulmonary Division, Heart Institute (InCor), Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo (HCFMUSP), Sao Paulo, SP, Brazil.

2 Radiology Institute (InRad), Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo (HCFMUSP), Sao Paulo, SP, Brazil.

3 Informatics Division, Heart Institute (InCor), Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo (HCFMUSP), Sao Paulo, SP, Brazil.

Correspondence to: Dr Carlos RR Carvalho (ORCID: 0000-0002-1618-8509), Pulmonary Division, Heart Institute (InCor), Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo (HCFMUSP), Av. Dr Eneas Carvalho de Aguiar, 44, Cerqueira Cesar, São Paulo, SP, 05403-900. Email: carlos.carvalho@hc.fm.usp.br. Fone: +55 11 26614505.

Word Count: 3550

Abstract

Objective This study aimed to propose a simple, accessible, and low-cost predictive clinical model to detect lung lesions due to COVID-19 infection.

Design This prospective cohort study included COVID-19 survivors hospitalised between March 30, 2020 and August 31, 2020 followed-up six months after hospital discharge. The pulmonary function was assessed using the modified Medical Research Council (mMRC) dyspnoea scale, oximetry (SpO_2), spirometry (forced vital capacity [FVC]), and chest X-ray (CXR) during an in-person consultation. Patients with abnormalities in at least one of these parameters underwent chest computed tomography (CT). mMRC scale, SpO2, FVC, and CXR findings were used to build a machine learning model for lung lesion detection on CT.

Setting A tertiary hospital in Sao Paulo, Brazil.

Participants 749 eligible RT-PCR-confirmed SARS-CoV-2 infected patients aged ≥18 years.

Primary outcome measure A predictive clinical model for lung lesion detection on chest CT.

Results There were 470 patients (63%) that had at least one sign of pulmonary involvement and were eligible for CT. Almost half of them (48%) had significant pulmonary abnormalities, including ground-glass opacities, parenchymal bands, reticulation, traction bronchiectasis, and architectural distortion. The machine learning model, including the results of 257 patients with complete data on mMRC, SpO_2, FVC, CXR and CT, accurately detected pulmonary lesions by the joint data of CXR, mMRC scale, SpO_2, and FVC (sensitivity, 0.85±0.08; specificity, 0.70±0.06; F1-score, 0.79±0.06; and AUC, 0.80±0.07).

Conclusion A predictive clinical model based on CXR, mMRC, oximetry, and spirometry data can accurately screen patients with lung lesions after SARS-CoV-2 infection. Given that these examinations are highly accessible and low cost, this protocol can be automated and implemented in different countries for early detection of COVID-19 sequelae.

Strengths and limitations of this study

- This study conducted a broad clinical assessment, embracing an in-person functional, and radiological pulmonary examinations of a large cohort of COVID-19 patients.

- The sample size used for artificial intelligence evaluation was sufficient to provide a robust prediction equation.

- Although the study was conducted in a single centre, the cohort population was heterogeneous and hailed from all districts of the metropolitan region of Sao Paulo (with approximately 21 million inhabitants).

- Although there were some missing patient data and data lost to follow-up, in general they were from patients that had less severe disease and were less likely to develop lung lesions.

INTRODUCTION

The coronavirus disease 2019 (COVID-19) caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) emerged in December 2019 and had since spread globally.[1] This multisystemic viral disease promotes endothelial and microvascular damage and immune system dysregulation, leading to hyperinflammatory and hypercoagulable states.[2 3] Several organs can be affected during the acute phase of COVID-19. In particular, pulmonary complications are considered life-threatening owing to the risk of progression to respiratory failure.[4 5]

COVID-19 symptoms can persist for more than 12 weeks after acute infection, characterizing long COVID.[1] The clinical complains of dyspnoea, fatigue, cough, chest pain, depression, cognitive disorders, headache, palpitations, myalgia, and arthralgia are the most reported in long COVID.[6-9] In addition to symptoms, some studies have shown that radiological abnormalities are also frequent in the follow-up of patients after the acute phase. One study performed chest computed tomography (CT) in 171 patients 4 months after hospital discharge and showed abnormalities in 75.5% of the patients who required invasive mechanical ventilation (IMV).[10] “Fibrotic-like changes” were observed in 19.3% of the total cohort and in 38.8% of patients with acute respiratory distress syndrome.[9] IMV can predict pulmonary sequelae, which reduce functional capacity and the health-related quality of life.[6 11 12] The National Institute for Health and Care Excellence (NICE), has reported that some examinations can guide the diagnosis and management of post-COVID-19 syndrome,[1] including oximetry, spirometry, chest X-ray (CXR), ultrasonography, modified Medical Research Council (mMRC) dyspnoea scale, and chest CT. The latter examination is the gold standard for the diagnosis of chronic lung lesions due to COVID-19 and characterization of “fibrotic-like” lung lesions.[1 10]

The World Health Organization reported more than 265 million confirmed COVID-19 cases worldwide, with approximately 5 million deaths, and 260 million

patients recovered as of December 2021.[13] The large number of recovered individuals experiencing long-term COVID-19 symptoms, such as fatigue, weakness, and dyspnoea, has drawn the attention of researches,[14 15] as they are expected to impose a significant health and economic burden.[14] In early 2021, the United Kingdom National Institute for Health Research invested £18.5 million to fund studies on long COVID.[16] The lack of knowledge and medical training for treating post-COVID symptoms also represents a significant public health challenge.[14] Thus, health care systems will have to reorganize themselves to address this issue, requiring the reallocation of resources and training of multidisciplinary teams and the development of new approaches.[14]

In this context, the wide availability of CXR and CT scanners has enabled the development of deep learning (DL) artificial intelligence-based algorithms for the automated diagnosis and prognosis of COVID-19.[17-19] For example, Castiglioni et al. [17] proposed a DL model for diagnosing COVID-19 with high sensitivity and specificity using radiography findings, whereas Wang et al. [18] developed a DL model (DenseNet) to classify CT images as positive or negative for COVID-19.

Although these studies presented promising results, they focused on images of patients in the acute phase of COVID-19. However, as the pandemic is still ongoing with limited knowledge on long COVID-19 consequences,[20] a more comprehensive protocol for screening COVID-19 patients and assessing the risk of chronic pulmonary changes in recovered patients has not been validated to date. Thus, this study aimed to develop a predictive clinical model to detect the presence of radiologic chronic lung lesions due to SARS-CoV-2 infections based on the results of simple and accessible examinations, such as the mMRC dyspnoea scale, oximetry, spirometry, and CXR.

METHODS

Study design and eligibility

This prospective cohort study detected lung lesions in adult patients (≥ 18 years) with RT-PCR-confirmed SARS-CoV-2 infection admitted to the ward or intensive care unit (ICU) of the Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo (HCFMUSP), Sao Paulo, Brazil, from March 30 to

August 31st, 2020. The RT-PCR-confirmed SARS-CoV-2 infection was obtained at hospital admission day. We considered only the first admission of each patient on the HCFMUSP. The protocols used in this study were described previously.[21] All research procedures were approved by the Research Ethics Committee of our institution (Process No. 31942020.0.000.0068).

The patients were invited to participate in the study six months after admission, and a face-to-face consultation was scheduled. At this point, all patients were already discharged. Clinical, radiological, and laboratory evaluations were performed at face-to-face consultations after the patients gave written informed consent. Clinical data (comorbidities, cardiorespiratory symptoms, and smoking history), including the length of ICU stay and the need for IMV, were retrospectively collected from the electronic medical records of HCFMUSP. All data were stored in a structured form developed using REDCap software (https://www.redcapbrasil.com.br/).

General evaluation

Patients who agreed to participate in the study signed an informed consent form and underwent a face-to-face consultation during the collection of anthropometric data and a pulmonary assessment, with an emphasis on respiratory symptoms. Dyspnoea was assessed using the mMRC scale.[21] Oxygen saturation (SpO_2) at rest and after physical exertion (1-min sit and stand test) was measured by pulse oximetry.[21 22] Spirometry was performed according to criteria established by ATS/ERS Task Force.[23] Actual spirometry results were compared with predicted values, according to Pereira et al. [24].

At the same face-to-face consultation described above, the same patients underwent a posteroanterior and lateral CXR according to standard guidelines. The results of these examinations were evaluated blindly and independently by two chest radiologists (MVYS and RCC, have 7 and 16 years of experience in thoracic radiology, respectively) working on dedicated workstations. The radiographs were scored as 0 (results were normal or not related to COVID-19 [including cardiomegaly and pulmonary nodules, for instance]) or 1 (findings which could be related to COVID-19 [including bilateral linear and/or reticular

opacities, especially peripheral opacities]). Disagreements were resolved by consensus. The agreement rate was 75%.

After the consensus classification performed by the radiologists (described above), the dataset with classified CXR were used to train and validate a DL algorithm developed to predict the probability that the CXR had findings related to sequelae of COVID-19. The DL algorithm is based on an EfficientNetB7 architecture[25] and a five-fold cross-validation strategy was adopted to train and validate the model, leading to an average area under the curve (AUC) of 0.89 (Supplemental Methods).

Chest CT

Patients who meet at least one the following criteria during the general evaluation were enrolled to undergo CT: (a) mMRC ≥ 2; (b) resting SpO_2 ≤ 90% and/or a decrease in SpO_2 of ≥4% during the 1-min sit and stand test; (c) opacities likely related to COVID-19 on CXR; and (d) FVC < lower limit of normal (LLN). The mean interval between CXR and chest CT was 45 ± 33 days.

The CT protocol used in this study was described previously.[21] CT findings consistent with COVID-19, including ground-glass and peripheral opacities, consolidations, parenchymal bands, reticulations, traction bronchiectasis, architectural distortions, honeycombing, bronchial wall thickening, mosaic attenuation, and pleural effusion, were categorized according to the criteria of the Fleischner Society.[26] The extent of lung involvement was quantified according to Francone et al. [27] by assigning the following scores to each pulmonary lobe: 0, none; 1, <5%; 2, 5-25%; 3, 26-50%; 4, 51-75%; and 5, >75%. The total score varied from 0 to 25 and was calculated by summing the scores of the five lobes. [25] Categorization of the CT features and score assignment were blindly and independently performed by the same two thoracic radiologists who evaluated the CXR (MVYS and RCC). Any disagreements were resolved by consensus.

A score ≥7 was used as the cut off value for significant CT changes after model calibration. The equations used to determine these scores are described in the Supplemental Methods.

Machine learning (ML) model

A Machine Learning (ML) model based on a Logistic Regression (LR) with L2 regularization to prevent overfitting[28] was adopted to detect the presence of COVID-19-related chronic lung lesions. The L1 regularization was not included due to the variable selection by statistical significance that removed irrelevant and correlated attributes. In this ML model, the results of the mMRC scale, oximetry, spirometry, and DL-based classification of 257 CXR images were used as input data, and the presence of pulmonary lesions was used as output data (Figure 1). The performance of the model was evaluated by the metrics sensitivity, specificity, AUC, and F1-score after a five-fold cross validation. (Supplemental Methods)

Statistical analysis

Continuous variables are expressed as the mean and standard deviations or median and interquartile range. Normality of the variables was assessed by D'agostino-Pearson test. Normally and non-normally distributed continuous variables were compared using the Student's *t*-test and Mann-Whitney U test, respectively. Categorical variables are presented as counts and percentages and compared using the chi-square test. (Excel 2016; Python 3.8.11; extension packages: Pandas 1.0.1; Numpy 1.19.5; Scipy 1.5.4; Scikit-Learn 0.24.0).

The performance of the DL models was assessed by the area under the receiver operating characteristic (AUC) curve. The performance of the ML model was determined based on sensitivity, specificity, F1-score and AUC values (Supplemental Methods).

Patient and public involvement

Patients or the public were not involved in the design, conduct, reporting or dissemination plans of this research.

RESULTS

Of 3,753 COVID-19 enrolled patients, 1,957 were eligible for the study and 749 were included in the final analysis (445 [59%] and 304 [41%] patients were admitted to the ICU and ward, respectively). Additional information on the inclusion and exclusion criteria is shown in Figure 2.

Demographic characteristics of the cohort are shown in Supplemental Table S1. The median age was 56 years, with a predominance of overweight individuals, and 53% were male. Additionally, 59.4% of the patients were admitted to the ICU, and 68.5% of them were on IMV during the study period. The vital signs of most patients were within normal limits during the hospitalisation period (Supplemental Table S1).

The median interval between hospital admission and consultation was 7.1 (IQR [6.7–8.5]) months, and the minimum and maximum values of this interval were 5.4 and 12.9 months, respectively. Of the 749 patients, 470 (63%) had at least one sign of pulmonary involvement (Table 1). Supplemental Figure S1 illustrates the simultaneous presence of two or more criteria for pulmonary involvement.

Table 1. Pulmonary function of patients with signs of pulmonary involvement (N=749).

Variables	Patients with signs of pulmonary involvement (N=749)
mMRC ≥ 2	229/742 (30.9)
Altered Oximetry*	71/675 (10.5)
CXR (score 1)	200/629 (31.8)
FVC < LLN	212/642 (33)

Values are presented as n/N (%). CXR, chest X-ray; FVC, forced vital capacity; mMRC, modified Medical Research Council dyspnoea scale; LLN, lower limit of normal. *Resting SpO_2 ≤90% or a decrease in SpO_2 of ≥4% during the 1-min sit and stand test.

The demographic and clinical characteristics of patients stratified by the presence of pulmonary involvement are described in Supplemental Table S2. Patients with pulmonary involvement were older and predominantly female, have more comorbidities, and a higher rate of ICU admission than those without (Supplemental Table S2). In patients with pulmonary involvement, 348 underwent CT (68%) (Figure 2). The demographic and clinical characteristics were similar between those that underwent or did not undergo the CT (Supplemental Table S3).

CT scores were obtained from 328 (94%) patients. Scores were not determined in 20 patients, who were excluded because of low CT scan quality or had motion artefacts. Chest CT analysis showed that 47.6% of the patients had a score ≥7, and the most common features were ground-glass opacities, parenchymal bands, reticulation, traction bronchiectasis, and architectural distortions (Supplemental Table S4). In this group, 86.5% and 13.5% were admitted to the ICU and ward, respectively. Among the patients with normal CT findings (score = 0), 36.4% and 63.6% were admitted to the ICU and ward, respectively. The frequency of CT changes is shown in Supplemental Table S5. That frequency of “fibrotic-like” lesions, including traction bronchiectasis and architectural distortion, was significantly higher in the group admitted to the ICU in the acute phase of the disease. Long-term CT features in patients with moderate and critical COVID-19 are shown in Figure 3 and Supplemental Figure S2, respectively.

Of the 348 patients with CT data, 257 had data on mMRC, oximetry, spirometry, X-ray, and chest CT and were selected for the prediction analysis of pulmonary changes. Among the 91 patients excluded for the prediction analysis, 61 had incomplete data of all four tests (mMRC, oximetry, spirometry, CXR and CT) and 30 showed radiographic signs not related to COVID-19 (Supplemental Table S6).

Three data groups were considered for the prediction analysis of pulmonary changes: (1) clinical data (oximetry [SpO_2], mMRC dyspnoea scores, and spirometry [FVC]), (2) CXR, and (3) all variables (oximetry [SpO_2], mMRC dyspnoea scores, spirometry [FVC], and CXR). The performance of the predictive

model was higher using the combination of all variables (clinical variables and CXR), and the following metrics expressed in terms of mean ± standard deviation and 95% Confidence Interval (CI) were observed: sensitivity, 0.85±0.08 (95% CI [0.77, 0.94]); specificity, 0.70±0.14 (95% CI [0.55, 0.85]); F1-score, 0.79±0.06 (95% CI [0.73, 0.85]); and AUC, 0.80±0.07(95% CI [0.72, 0.87]) (Table 2).

Table 2. Performance of the predictive model using three combinations of variables (N=257).

Groups of variables	Sensitivity	Specificity	F1-score	AUC
1 SpO_2, mMRC score, and FVC	0.87±0.16	0.42±0.33	0.71±0.03	0.68±0.10
2 CXR	0.88±0.05	0.52±0.14	0.75±0.04	0.78±0.05
3 SpO_2, mMRC score, FVC, and CXR	0.85±0.08	0.70±0.14	0.79±0.06	0.80±0.07

Values are presented as means ± standard deviations after five-fold cross validation for each test fold. CXR, chest X-Ray; FVC, forced vital capacity; mMRC, Modified Medical Research Council dyspnoea scale.

The machine learning predictive model is represented by the following function:

$$p_{CT} = \sigma\left(\beta_1 FVC^* + \beta_2\, mMRC^* + \beta_3 S_pO_2 + \beta_4 p_{CXR0} + \beta_5 p_{CXR1} + \beta_6 p_{CXR2} + \beta_7 p_{CXR3} + \beta_8 p_{CXR4}\right)$$

$$\beta_1 = -0.3705 \;\; \beta_2 = -2.2807 \;\; \beta_3 = -0.745 \;\; \beta_4 = 1.1257$$

$$\beta_5 = 1.4960 \;\; \beta_6 = 1.0761 \;\; \beta_7 = 0.7328 \;\; \beta_8 = -0.7613$$

Where p_{CT} is the probability of the presence of abnormalities on CT images, σ is the sigmoid function to restrict p_{CT} between 0 and 1, $FVC^* = \frac{FVC_{Resting}}{2FVC_{min}}$, $mMRC^* = \frac{mMRC}{4}$, and p_{CXR0} to p_{CXR4} are the probabilities that the CXR image has findings related to sequelae from COVID-19, obtained in each fold (0 to 4) during a 5-folds cross validation. (Supplemental Methods)

Therefore, based in these observations, we propose in a flowchart a suggestion for lung lesion case-finding in COVID-19 survivors (Figure 4).

DISCUSSION

Few studies have assessed the pulmonary abnormalities in COVID-19 survivors after six months of hospital discharge. However, some of these patients have developed long-term pulmonary complications after the acute phase of the disease.[6 29-33] This study evaluated 749 COVID-19 patients who received supplemental oxygen or ventilatory support in the ward or ICU and survived. They underwent an in-person comprehensive clinical, functional, and radiological assessments, which were more extensive than those performed in previous studies,[6 30 31 33-35] conferring reliability to our results.

In the first months after recovery, the most common CT findings in COVID-19 hospitalised patients included ground-glass opacities, parenchymal bands, reticulation, mosaic attenuation pattern, and "fibrotic-like" abnormalities, including traction bronchiectasis and architectural distortions.[36 37] These findings were detected in 76.5% of our cohort, and severe and extensive changes were noted in approximately 50% of the cases. The CT abnormalities were more prevalent in older critical patients and individuals with more comorbidities, which is consistent with previous studies.[32 38] These results indicate the high prevalence of chronic lung lesions and sequelae in post-COVID patients worldwide.

Therefore, the need to identify severe pulmonary complications due to COVID-19, including fibrosis,[1] and the large number of COVID-19 survivors, prompted us to develop a predictive clinical model to screen patients admitted to a tertiary hospital, which could be able to reduce costs and radiation exposure. During the first six months of the pandemic in Sao Paulo, Brazil, all hospital beds at HCFMUSP (300 in the ICU and 400 in the ward) were made available to COVID-19 patients.[12] Patients were treated free of charge in our hospital owing to a universal health system, and there is a constant search for better and cost-effective protocols to improve workflow.[12]

Dyspnoea scales, CXR, oximetry, and spirometry are commonly used to evaluate COVID-19 symptoms.[2] A Norwegian study evaluated a cohort of 100 patients three months after admission to a hospital and reported that 19% had dyspnoea (mMRC score>1) and 10% presented altered FVC and normal oxygen saturation levels, suggesting the lower sensitivity of pulse oximetry.[39] In 113 patients evaluated 4 months after COVID-19 diagnosis in Switzerland, FVC and oxygen saturation levels were lower in patients who had a severe disease than in those with a moderate disease, although the mean values remained within the limits of normality.[35] In addition, a previous study has suggested that cough, lymphocytosis and the lung volume could indicate lung lesions in COVID-19-recovered patients.[34]

Ground-glass and reticular opacities can be detected by CXR, although this method is less sensitive than CT.[40] On the other hand, CXR is readily available in the primary care setting and has a lower cost and radiation exposure than CT.[40 41] Radiographs were separately scored by an automated DL-based image analysis tool and chest radiology specialists, and there was a high level of consensus between these scores (AUC = 0.89). In the Brazilian public health system, the cost of a CT scan is approximately 15 times higher than that of a CXR.[41] According to the American College of Radiology and the Radiological Society of North American, the radiation doses of a standard chest CT and CXR are 6.1 mSv and 0.1 mSv, respectively; this underscores the advantage of CXR in reducing the exposure of COVID-19 patients to radiation, especially those who have already performed serial imaging exams in the acute phase of the disease.[42]

Nevertheless, none of these examinations alone accurately predicted pulmonary complications. The performance of our model corroborates this finding since the information provided by each clinical examination alone did not accurately diagnose the pulmonary changes detected on CT. In contrast, clinical and radiographic data were complementary and increased the performance of the ML model. Cross-validation also increased the robustness of the results. These results indicate that four examinations (oximetry, mMRC dyspnoea scale, spirometry, and CXR) should be jointly conducted to screen patients at risk of developing chronic lung lesions due to COVID-19 and achieve a diagnostic performance similar to that of CT (sensitivity, 0.85±0.08; specificity, 0.70±0.14;

F1-score, 0.79±0.06; and AUC, 0.80±0.07). Analysis of these metrics indicates that this predictive clinical method can better identify the true positives than true negatives. In addition, the F1-score takes into account both false-positive and false-negative results and measures the accuracy of the method in the dataset.

The WHO has highlighted the importance of establishing screening protocols with a favourable cost-effectiveness ratio for patients affected by different pathologies.[43] The identification of COVID-19 lung lesions will allow the accurate referral of patients to specialists for further investigation and treatment. As the COVID-19 sequelae can progress to increasing intensity of symptoms and risk of disability, this approach can improve the quality and length of life of patients, since medical interventions can be performed as early as possible.

We already have an initiative to implement this protocol in Brazil. The project will start in the state of Sao Paulo, in partnership with the State of Sao Paulo Health Department, where the HCFMUSP is located. We will start to apply this screening protocol in the central area of the city of Sao Paulo, with approximately 430,000 inhabitants, according to the flowchart suggested for lung lesion case-finding in COVID-19 survivors (Figure 4). Firstly, exams will be performed in the following order, starting from the simplest and most accessible ones: oximetry/mMRC, spirometry and CXR. At the moment the patient shows alterations in any of these four exams, the patient will be enrolled directly for further investigation in a specialised care centre to perform CT and/or other specific exams. We expect that over time, this can lead to a significant reduction in morbidity and mortality due to COVID-19 lung sequelae, relieving the burden on the health care system, reducing expenses of imaging exams and accelerating the medical interventions.

This study has some limitations. First, there was variability in the interval between the execution of CXR and CT. Notwithstanding this variation, which might contribute to lung recovery, our protocol screened a large number of patients with pulmonary lesions, demonstrating the persistence of these manifestations secondary to COVID-19 and reducing sampling bias. Second, the single-centre nature of the study limits the generalizability of our results. However, a previous study showed that the population of patients admitted to

HCFMUSP—a tertiary reference hospital for the treatment of COVID-19 in Brazil—was heterogeneous and hailed from all districts of the metropolitan region of Sao Paulo (with approximately 21 million inhabitants).[12] Third, we were unable to contact some patients because of inconsistencies in telephone numbers and addresses. Thus, these subjects were not included in the protocol, although public death registry data showed that they were alive. Fourth, this screening protocol was developed based on respiratory complaints, which are considered risk factors for the development of chronic lung complications. However, other COVID-19 symptoms were not analysed in this study.

The breadth of our results allowed us to propose a simple, accessible, and low-cost clinical predictive model to screen patients at risk of developing chronic lung lesions due to COVID-19. The low cost and easy accessibility to these examinations facilitate the implementation of the proposed protocol in developing countries. In addition, it may contribute to early and effective determination of the treatment course, thus reducing radiation exposure and the conduct of costly imaging examinations. The use of artificial intelligence facilitated the large-scale assessment of radiographs and their association with clinical variables, demonstrating that artificial intelligence models can be used to automate diagnosis, especially in severe patients.

Collaborators: *Members of the HCFMUSP Covid-19 Study Group: Adriana L Araújo, Aluisio C Segurado, Amanda C Montal, Anna Miethke-Morais, Anna S Levin, Beatriz Perondi, Bruno F Guedes, Carolina Carmo, Carolina S Lázari, Cassiano C Antonio, Clarice Tanaka, Claudia C Leite, Cristiano Gomes, Edivaldo M Utiyama, Emmanuel A Burdmann, Eloisa Bonfá, Esper G Kallas, Ester Sabino, Euripedes C Miguel, Fabio R Pinna, Fabiane Y O Kawano, Geraldo F Busatto, Giovanni G Cerri, Guilherme Fonseca, Heraldo P Souza, Izabel Marcilio, Izabel C Rios, Jorge Hallak, José Eduardo Krieger, Juliana C Ferreira, Julio F M Marchini, Larissa S Oliveira, Leila Harima, Linamara R Batisttella, Luis Yu, Luiz Henrique M Castro, Marcelo C Rocha , Marcello M C Magri, Marcio Mancini, Maria Amélia de Jesus, Maria Cassia J M Corrêa, Maria Cristina P B Francisco, Maria Elizabeth Rossi, Marjorie F Silva, Marta Imamura, Maura S Oliveira, Nelson Gouveia, Orestes V Forlenza, Paulo A Lotufo, Ricardo F Bento, Ricardo Nitrini, Rodolfo F Damiano, Roger Chammas, Rossana P Francisco, Solange R G

Fusco, Tarcisio E P Barros-Filho, Thais Mauad, Thaís Guimarães, Thiago Avelino-Silva and Wilson J Filho.

Funding: This study was funded by the Sao Paulo Research Foundation (grant number - 2020/07200-9).

Contributors: CRRC: conceptualisation, data curation, formal analysis, funding acquisition, investigation, methodology, project administration, resources, supervision, validation, visualisation, writing – original draft, and writing – review & editing. RCC: data curation, formal analysis, investigation, methodology, validation, visualisation, writing – original draft, and writing – review & editing. MVYS: data curation, formal analysis, investigation, methodology, validation, visualisation, writing – original draft, and writing – review & editing. MLG: conceptualisation, data curation, formal analysis, investigation, methodology, project administration, supervision, validation, visualisation, writing – original draft, and writing – review & editing. CAL: data curation, formal analysis, investigation, writing – original draft, and writing – review & editing. DACC: data curation, formal analysis, methodology, software, writing – original draft, and writing – review & editing. DML: data curation, formal analysis, methodology, software, writing – original draft, and writing – review & editing. PGS: conceptualisation, project administration, supervision, validation, visualisation and writing – review & editing. JMS: methodology, validation, visualisation and writing – review & editing. CHN: methodology, validation, visualisation and writing – review & editing. MAG: data curation, formal analysis, funding acquisition, methodology, software, supervision, validation, visualisation, writing – original draft, and writing – review & editing. HCFMUSP Covid-19 Study Group: contributed to the implementation of the study and data collection. All authors critically reviewed and approved the final version.

Competing interests' statement: None declared.

Data Sharing: The study protocol was previously described by Busatto et al. [21] and was registered at the "Brazilian Registry of Clinical Trials" (https://ensaiosclinicos.gov.br/). The raw data are not publicly available because follow-up studies will be carried out. However, data are available from the

corresponding author upon request and authorization from the institution. Data on demographics, hospitalisation, and outcomes are available in the COVID-19 Data Sharing/BR repository and are freely available for download[44].

Ethical approval: The study was approved by the Research Ethics Committee of the HCFMUSP (approval number 31942020.0.000.0068).

REFERENCES

1. Sisó-Almirall A, Brito-Zerón P, Conangla Ferrín L, et al. Long Covid-19: Proposed Primary Care Clinical Guidelines for Diagnosis and Disease Management. *Int J Environ Res Public Health* 2021;18(8) doi: 10.3390/ijerph18084350 [published Online First: 2021/04/20]
2. Nalbandian A, Sehgal K, Gupta A, et al. Post-acute COVID-19 syndrome. *Nat Med* 2021;27(4):601-15. doi: 10.1038/s41591-021-01283-z [published Online First: 2021/03/22]
3. Mauad T, Duarte-Neto AN, da Silva LFF, et al. Tracking the time course of pathological patterns of lung injury in severe COVID-19. *Respir Res* 2021;22(1):32. doi: 10.1186/s12931-021-01628-9 [published Online First: 20210129]
4. Tanni SE, Fabro AT, de Albuquerque A, et al. Pulmonary fibrosis secondary to COVID-19: a narrative review. *Expert Rev Respir Med* 2021;15(6):791-803. doi: 10.1080/17476348.2021.1916472 [published Online First: 2021/04/27]
5. Macedo BR, Garcia MVF, Garcia ML, et al. Implementation of Tele-ICU during the COVID-19 pandemic. *J Bras Pneumol* 2021;47(2):e20200545. doi: 10.36416/1806-3756/e20200545 [published Online First: 20210430]
6. Huang C, Huang L, Wang Y, et al. 6-month consequences of COVID-19 in patients discharged from hospital: a cohort study. *Lancet* 2021;397(10270):220-32. doi: 10.1016/S0140-6736(20)32656-8 [published Online First: 2021/01/08]
7. Lopez-Leon S, Wegman-Ostrosky T, Perelman C, et al. More Than 50 Long-Term Effects of COVID-19: A Systematic Review and Meta-Analysis. *Res Sq* 2021 doi: 10.21203/rs.3.rs-266574/v1 [published Online First: 2021/03/01]
8. Fernandes PMP, Mariani AW. Life post-COVID-19: symptoms and chronic complications. *Sao Paulo Med J* 2021;139(1):1-2. doi: 10.1590/1516-3180.2021.139104022021
9. Morin L, Savale L, Pham T, et al. Four-Month Clinical Status of a Cohort of Patients After Hospitalization for COVID-19. *JAMA* 2021;325(15):1525-34. doi: 10.1001/jama.2021.3331
10. Raghu G, Collard HR, Egan JJ, et al. An official ATS/ERS/JRS/ALAT statement: idiopathic pulmonary fibrosis: evidence-based guidelines for diagnosis and management. *Am J Respir Crit Care Med* 2011;183(6):788-824. doi: 10.1164/rccm.2009-040GL
11. Carfì A, Bernabei R, Landi F, et al. Persistent Symptoms in Patients After Acute COVID-19. *JAMA* 2020;324(6):603-05. doi: 10.1001/jama.2020.12603
12. Ferreira JC, Ho YL, Besen BAMP, et al. Protective ventilation and outcomes of critically ill patients with COVID-19: a cohort study. *Ann Intensive Care* 2021;11(1):92. doi: 10.1186/s13613-021-00882-w [published Online First: 2021/06/07]
13. WHO. WHO Coronavirus Disease (COVID-19) Dashboard 2021 [cited 2021 October 13]. Available from: https://covid19.who.int/ accessed October 13 2021
14. Wade DT. Rehabilitation after COVID-19: an evidence-based approach. *Clin Med (Lond)* 2020;20(4):359-65. doi: 10.7861/clinmed.2020-0353 [published Online First: 2020/06/09]

15. Godoy CG, Silva ECGE, Oliveira DB, et al. Protocol for Functional Assessment of Adults and Older Adults after Hospitalization for COVID-19. *Clinics (Sao Paulo)* 2021;76:e3030. doi: 10.6061/clinics/2021/e3030 [published Online First: 20210614]
16. Subbaraman N. US health agency will invest $1 billion to investigate 'long COVID'. *Nature* 2021;591(7850):356. doi: 10.1038/d41586-021-00586-y
17. Castiglioni I, Ippolito D, Interlenghi M, et al. Machine learning applied on chest x-ray can aid in the diagnosis of COVID-19: a first experience from Lombardy, Italy. *Eur Radiol Exp* 2021;5(1):7. doi: 10.1186/s41747-020-00203-z [published Online First: 20210202]
18. Wang S, Zha Y, Li W, et al. A fully automatic deep learning system for COVID-19 diagnostic and prognostic analysis. *Eur Respir J* 2020;56(2) doi: 10.1183/13993003.00775-2020 [published Online First: 20200806]
19. Ferreira Junior JR, Cardona Cardenas DA, Moreno RA, et al. Novel Chest Radiographic Biomarkers for COVID-19 Using Radiomic Features Associated with Diagnostics and Outcomes. *J Digit Imaging* 2021;34(2):297-307. doi: 10.1007/s10278-021-00421-w [published Online First: 20210218]
20. Greenhalgh T, Knight M, A'Court C, et al. Management of post-acute covid-19 in primary care. *BMJ* 2020;370:m3026. doi: 10.1136/bmj.m3026 [published Online First: 2020/08/11]
21. Busatto GF, de Araújo AL, Duarte AJDS, et al. Post-acute sequelae of SARS-CoV-2 infection (PASC): a protocol for a multidisciplinary prospective observational evaluation of a cohort of patients surviving hospitalisation in Sao Paulo, Brazil. *BMJ Open* 2021;11(6):e051706. doi: 10.1136/bmjopen-2021-051706 [published Online First: 2021/06/30]
22. van den Borst B, Peters JB, Brink M, et al. Comprehensive health assessment three months after recovery from acute COVID-19. *Clin Infect Dis* 2020 doi: 10.1093/cid/ciaa1750 [published Online First: 2020/11/21]
23. Miller MR, Hankinson J, Brusasco V, et al. Standardisation of spirometry. *Eur Respir J* 2005;26(2):319-38. doi: 10.1183/09031936.05.00034805
24. Pereira CA, Sato T, Rodrigues SC. New reference values for forced spirometry in white adults in Brazil. *J Bras Pneumol* 2007;33(4):397-406. doi: 10.1590/s1806-37132007000400008
25. Tan M, Le Q. Efficientnet: Rethinking model scaling for convolutional neural networks. *International Conference on Machine Learning* 2019:6105-14.
26. Hansell DM, Bankier AA, MacMahon H, et al. Fleischner Society: glossary of terms for thoracic imaging. *Radiology* 2008;246(3):697-722. doi: 10.1148/radiol.2462070712 [published Online First: 20080114]
27. Francone M, Iafrate F, Masci GM, et al. Chest CT score in COVID-19 patients: correlation with disease severity and short-term prognosis. *Eur Radiol* 2020;30(12):6808-17. doi: 10.1007/s00330-020-07033-y [published Online First: 2020/07/04]
28. Vittinghoff E, Glidden DV, Shiboski SC, et al. Regression Methods in Biostatistics: Linear, Logistic, Survival, and Repeated Measures Models. 2nd ed. New York: Springer-Verlag, 2012:1272.
29. Wu Q, Zhong L, Li H, et al. A Follow-Up Study of Lung Function and Chest Computed Tomography at 6 Months after Discharge in Patients with Coronavirus Disease 2019. *Can Respir J* 2021;2021:6692409. doi: 10.1155/2021/6692409 [published Online First: 20210213]
30. Huang L, Yao Q, Gu X, et al. 1-year outcomes in hospital survivors with COVID-19: a longitudinal cohort study. *Lancet* 2021;398(10302):747-58. doi: 10.1016/S0140-6736(21)01755-4
31. Han X, Fan Y, Alwalid O, et al. Fibrotic Interstitial Lung Abnormalities at 1-year Follow-up CT after Severe COVID-19. *Radiology* 2021:210972. doi: 10.1148/radiol.2021210972 [published Online First: 20210727]

32. Han X, Fan Y, Alwalid O, et al. Six-month Follow-up Chest CT Findings after Severe COVID-19 Pneumonia. *Radiology* 2021;299(1):E177-E86. doi: 10.1148/radiol.2021203153 [published Online First: 20210126]
33. Stylemans D, Smet J, Hanon S, et al. Evolution of lung function and chest CT 6 months after COVID-19 pneumonia: Real-life data from a Belgian University Hospital. *Respir Med* 2021;182:106421. doi: 10.1016/j.rmed.2021.106421 [published Online First: 20210418]
34. Caruso D, Guido G, Zerunian M, et al. Post-Acute Sequelae of COVID-19 Pneumonia: Six-month Chest CT Follow-up. *Radiology* 2021;301(2):E396-E405. doi: 10.1148/radiol.2021210834 [published Online First: 20210727]
35. Guler SA, Ebner L, Aubry-Beigelman C, et al. Pulmonary function and radiological features 4 months after COVID-19: first results from the national prospective observational Swiss COVID-19 lung study. *Eur Respir J* 2021;57(4) doi: 10.1183/13993003.03690-2020 [published Online First: 2021/04/29]
36. Liu C, Ye L, Xia R, et al. Chest Computed Tomography and Clinical Follow-Up of Discharged Patients with COVID-19 in Wenzhou City, Zhejiang, China. *Ann Am Thorac Soc* 2020;17(10):1231-37. doi: 10.1513/AnnalsATS.202004-324OC
37. Tabatabaei SMH, Rajebi H, Moghaddas F, et al. Chest CT in COVID-19 pneumonia: what are the findings in mid-term follow-up? *Emerg Radiol* 2020;27(6):711-19. doi: 10.1007/s10140-020-01869-z [published Online First: 20201109]
38. Solomon JJ, Heyman B, Ko JP, et al. CT of Postacute Lung Complications of COVID-19. *Radiology* 2021:211396. doi: 10.1148/radiol.2021211396 [published Online First: 20210810]
39. Sonnweber T, Sahanic S, Pizzini A, et al. Cardiopulmonary recovery after COVID-19: an observational prospective multicentre trial. *Eur Respir J* 2021;57(4) doi: 10.1183/13993003.03481-2020 [published Online First: 2021/04/29]
40. Jacobi A, Chung M, Bernheim A, et al. Portable chest X-ray in coronavirus disease-19 (COVID-19): A pictorial review. *Clin Imaging* 2020;64:35-42. doi: 10.1016/j.clinimag.2020.04.001 [published Online First: 20200408]
41. DataSUS. MdS. SIGTAP - Sistema de Gerenciamento da Tabela de Procedimentos, Medicamentos e OPM do SUS. 2021 [cited 2021 January 03]. Available from: http://sigtap.datasus.gov.br/tabela-unificada/app/sec/inicio.jsp accessed January 03 2021.
42. Mettler FAM, Mahadevappa Bhargavan, MythreyiChambers, Charles E.Elee, Jennifer G. Frush, Donald P. Milano, Michael T. Miller, Donald L. Royal, Henry D. Spelic, David C.Ansari, Armin J. Bolch, Wesley E.Guebert, Gary M. Sherrier, Robert H. Smith, James M. Vetter, Richard J. Report Na. 184 - Medical Radiation Exposure of Patients in the United States. United States: National Council on Radiation Protection and Measurements, 2019.
43. Wilson J, Jungner G. Principles and Practice of Screening for Disease.: World Health Organization Public Health Papers, 1968.
44. FAPESP. COVID-19 DataSharing/BR 2021 [cited 2021 September 08]. Available from: https://repositoriodatasharingfapesp.uspdigital.usp.br/.

Figure Legends

Figure 1. Logistic regression-based machine learning model to detect the presence of COVID-19-related lung lesions. The patients were invited to participate in the study six months after COVID-19 positive RT-PCR at hospital admission. The modified Medical Research Council (mMRC) dyspnoea scale, oximetry (SpO2), spirometry (forced vital capacity [FVC]), and the five radiographic scores obtained during DL-based classification of CXR (pCXR) were used as input data, and the presence of lung lesions due to COVID-19 was used as output data. AI: artificial intelligence. CT: computed tomography.

Figure 2. Flowchart of patient selection. FVC, forced vital capacity; LLN, lower limit of normal; mMRC, modified Medical Research Council dyspnoea scale. *Rest SpO2 < 90% or a decrease in SpO2 of at least 4% after the 1-min sit and stand test.

Figure 3. Fibrotic-like changes after critical COVID-19 in a patient in his early 70s. (A) PA chest radiograph obtained 7 months after infection shows reticular opacities with a slight peripheral predominance diffusely distributed in both lungs. (B) Image from the same radiograph analysed by the AI algorithm with a heat map highlighting the areas of pulmonary involvement. (C, D) Chest CT obtained 8 months after infection shows moderate ground glass opacities, linear multifocal and reticular abnormalities, discrete traction bronchiectasis and slight parenchymal architectural distortion. The patient had dyspnoea (mMRC=1) and altered FVC (2.34 L / 60% pred), besides the normal oximetry (97%).

Figure 4. Flowchart for lung lesion case-finding in COVID-19 survivors. *Altered oximetry: Resting SpO2 ≤90% or a decrease in SpO2 of ≥4% during the 1-min sit and stand test. **Altered CXR: COVID-19 findings, including bilateral linear and/or reticular opacities, especially peripheral opacities. † The in-person consultation also should start with oximetry and mMRC examinations. †† The suggestion is to perform plethysmography with diffusion capacity measure. CXR, chest X-Ray; FVC, forced vital capacity; LLN, lower limit of normal; mMRC, modified Medical Research Council dyspnoea scale.

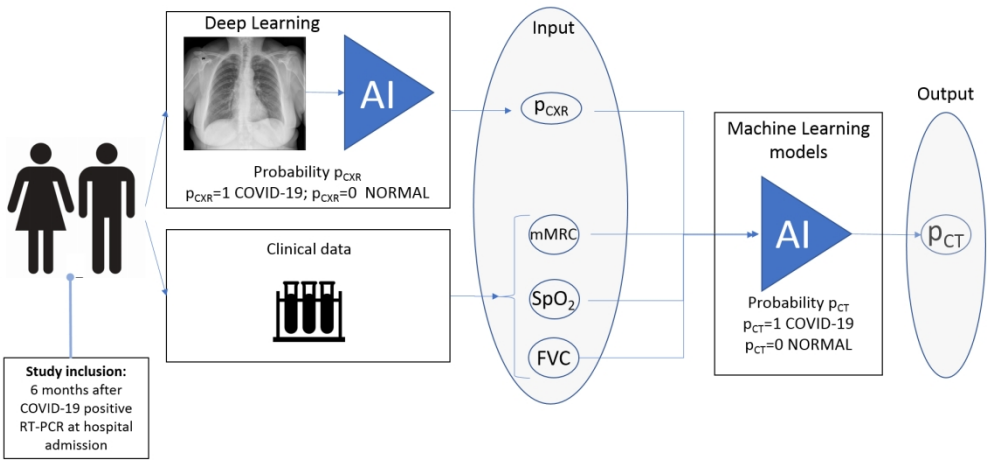


Figure 1. Logistic regression-based machine learning model to detect the presence of COVID-19-related lung lesions. The patients were invited to participate in the study six months after COVID-19 positive RT-PCR at hospital admission. The modified Medical Research Council (mMRC) dyspnoea scale, oximetry (SpO2), spirometry (forced vital capacity [FVC]), and the five radiographic scores obtained during DL-based classification of CXR (pCXR) were used as input data, and the presence of lung lesions due to COVID-19 was used as output data. AI: artificial intelligence. CT: computed tomography.

80x37mm (600 x 600 DPI)

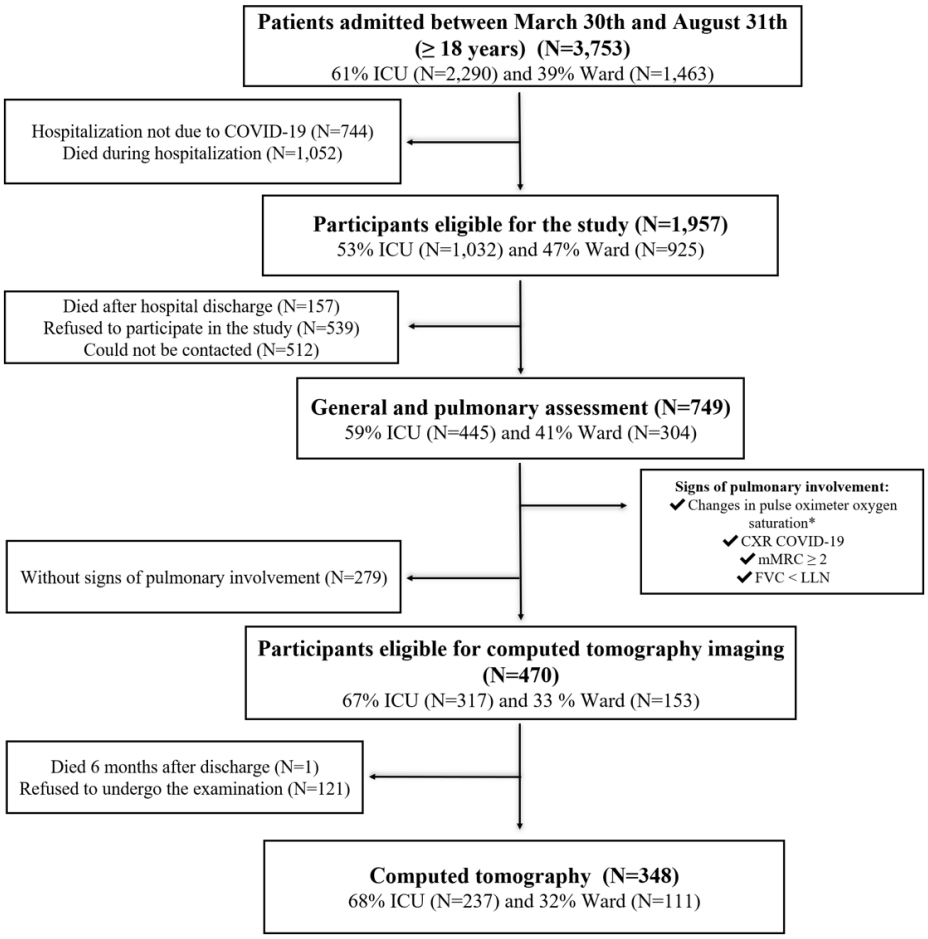


Figure 2. Flowchart of patient selection. FVC, forced vital capacity; LLN, lower limit of normal; mMRC, modified Medical Research Council dyspnoea scale. *Rest SpO2 < 90% or a decrease in SpO2 of at least 4% after the 1-min sit and stand test.

51x47mm (600 x 600 DPI)

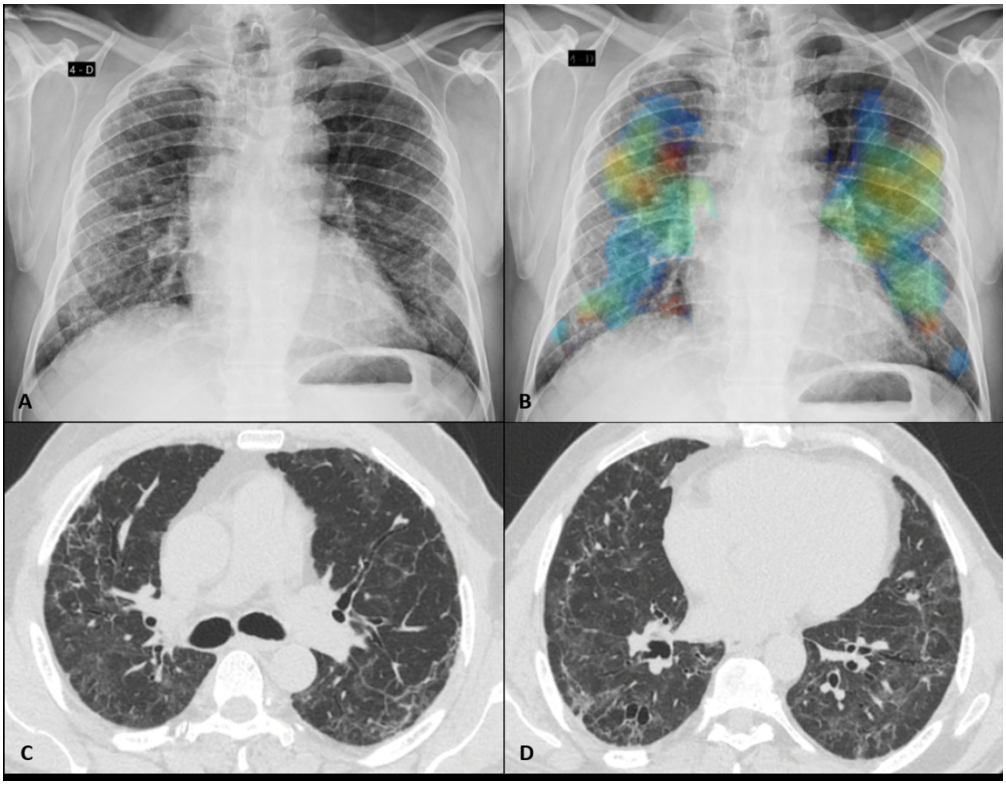


Figure 3. Fibrotic-like changes after critical COVID-19 in a patient in his early 70s. (A) PA chest radiograph obtained 7 months after infection shows reticular opacities with a slight peripheral predominance diffusely distributed in both lungs. (B) Image from the same radiograph analysed by the AI algorithm with a heat map highlighting the areas of pulmonary involvement. (C, D) Chest CT obtained 8 months after infection shows moderate ground glass opacities, linear multifocal and reticular abnormalities, discrete traction bronchiectasis and slight parenchymal architectural distortion. The patient had dyspnoea (mMRC=1) and altered FVC (2.34 L / 60% pred), besides the normal oximetry (97%).

154x119mm (600 x 600 DPI)

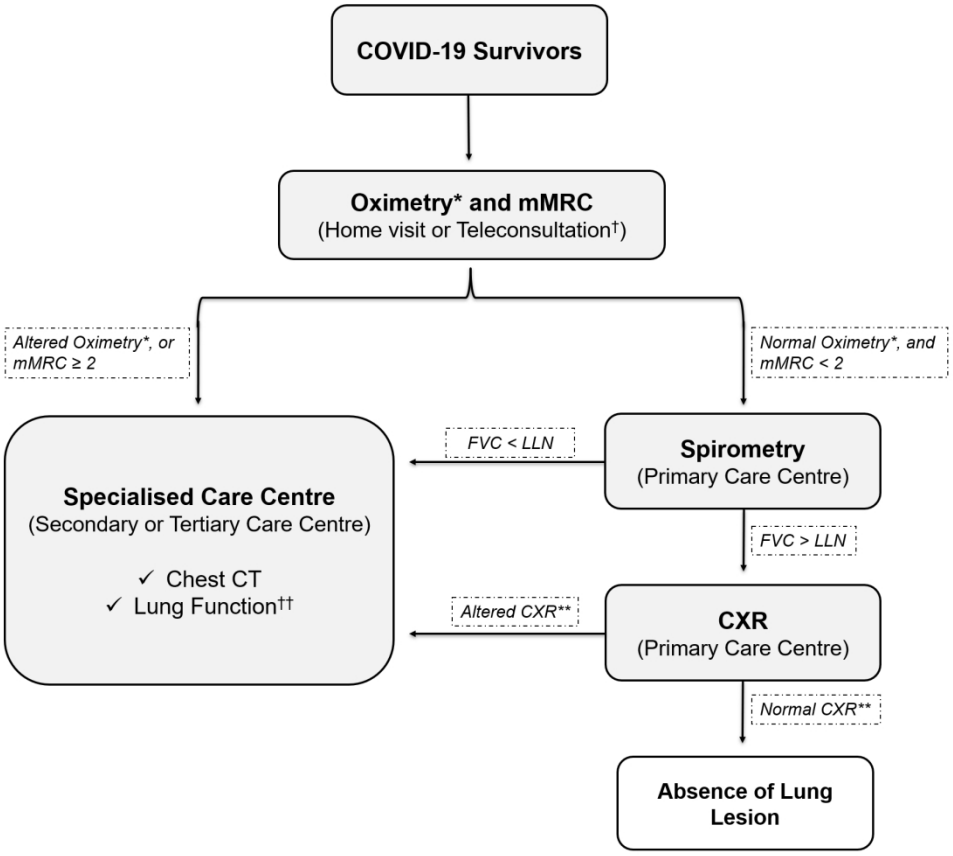


Figure 4. Flowchart for lung lesion case-finding in COVID-19 survivors. *Altered oximetry: Resting SpO2 ≤90% or a decrease in SpO2 of ≥4% during the 1-min sit and stand test. **Altered CXR: COVID-19 findings, including bilateral linear and/or reticular opacities, especially peripheral opacities. † The in-person consultation also should start with oximetry and mMRC examinations. †† The suggestion is to perform plethysmography with diffusion capacity measure. CXR, chest X-Ray; FVC, forced vital capacity; LLN, lower limit of normal; mMRC, modified Medical Research Council dyspnea scale.

81x71mm (600 x 600 DPI)

Supplemental Material

Chronic lung lesions in COVID-19 survivors: predictive clinical model

Carlos R R Carvalho, Rodrigo C Chate, Marcio VY Sawamura, Michelle L Garcia, Celina A Lamas, Diego AC Cardenas, Daniel M Lima, Paula G Scudeller, João M Salge, Cesar H Nomura, Marco A Gutierrez, HCFMUSP Covid-19 Study Group.

Contents

Supplemental Methods

Datasets

The SIIM-RSNA dataset contains 6,334 posterior-anterior radiographic images from 6,054 patients obtained from the public dataset Machine Learning Challenge on COVID-19 Pneumonia Detection and Localization.[1] Specialists classified images as “negative for pneumonia” or “COVID-19 pneumonia”. A total of 6,030 images were selected and randomly distributed in training and validation sets (1,276 negative and 3,711 positive findings) and a test set (400 negative and 643 positive findings).

The Institute of Radiology (InRad) dataset contains chest X-Ray (CXR) and chest computed tomographic (CT) images of 257 patients. The CXR images were classified as normal (n=145) or with findings related to COVID-19 (n=112) and randomly distributed in training and validation sets (214 patients) and a test set (n=43). Images were obtained from the InRad of the Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo (HCFMUSP).

Because of differences in dataset sizes, a data augmentation technique was adopted using random transformations, including rotation (0–15 degrees), horizontal mirroring, and random changes in intensity and contrast (0–5%).

Classification of chest radiography images

A deep-learning (DL) approach using a convolutional neural network (CNN) based on an EfficientNetB7 architecture was used.[2] The network classification layer was replaced by a global average pooling operation, followed by batch normalization and the adoption of a dense layer with one neuron and sigmoid activation function. Each training iteration was run for 40 epochs with an Adam optimizer at a learning rate of 0.0001. All images were resized to 600 x 600 pixels.

The CNN was trained using the SIIM-RSNA dataset to detect radiographic patterns of COVID-19 pneumonia. Training was initiated in EfficientNetB7 using weights after pre-training with the ImageNet dataset.[3]

A five-fold cross-validation strategy was adopted for the training and validation sets. The training weights obtained for each fold were used with the

test set of the SIIM-SNA to evaluate classification accuracy (Table 1). The fold with the best area under the receiver operating characteristic curve (AUC), in this case, fold 1 with AUC of 0.89, defines the final weights of the CNN.

Table 1. Classification of the test set of the SIIM-RSNA dataset as negative (normal) or positive (patterns of COVID-19 pneumonia).

Dataset	5-fold	Acc	Prec	Sensitivity	Specificity	F1-score	AUC
SIIM-RSNA	0	0.80	0.85	0.82	0.76	0.83	0.88
	1	0.80	0.85	0.82	0.77	0.84	**0.89**
	2	0.78	0.77	0.92	0.56	0.84	0.87
	3	0.76	0.74	0.93	0.48	0.83	0.86
	4	0.76	0.74	0.93	0.48	0.83	0.86

Area under the receiver operating characteristic curve (AUC); Accuracy (Acc); Precision (Prec).

For the InRad dataset, the CNN was initialized with the final weights defined in the training set of SIIM-RSNA. After initialization, the CNN was retrained to classify images as normal or with findings related to COVID-19.

The InRad dataset was divided into six-folds during the retraining, five folds for training and validation, and one-fold for test. To avoid bias, the test fold was selected to run all six folds available and, for each test fold selected, a five-fold cross-validation strategy was applied in the remaining training and validation folds (Table 2).

Table 2. Classification using six test folds of the InRad database.

Dataset	Test fold	Acc	Prec	Sensitivity	Specificity	F1-score	AUC
InRad	0	0.79±0.01	0.74±0.04	0.82±0.07	0.77±0.06	0.78±0.02	0.86±0.02
	1	0.69±0.02	0.62±0.03	0.84±0.06	0.57±0.07	0.71±0.02	0.75±0.01
	2	0.67±0.05	0.60±0.06	0.81±0.08	0.57±0.13	0.68±0.02	0.76±0.02
	3	0.77±0.04	0.71±0.07	0.80±0.04	0.74±0.10	0.75±0.03	0.80±0.02
	4	0.82±0.05	0.77±0.11	0.89±0.10	0.78±0.14	0.81±0.03	0.89±0.04
	5	0.71±0.04	0.62±0.04	0.90±0.02	0.58±0.08	0.73±0.03	0.80±0.02

Data represent the mean and standard deviation after five-fold cross validation. Area under the receiver operating characteristic curve (AUC); Accuracy (Acc); Precision (Prec).

Detection of chronic lung lesions on computed tomography images

Three machine learning models were developed based on the clinical data, including the modified Medical Research Council dyspnea scale (mMRC), oximetry (SpO_2) and spirometry (forced vital capacity, FVC), and five radiographic

probabilities (p_{CXR0} top_{CXR4}) with findings related to COVID-19 (p_{CXRn}=1) and normal (p_{CXRn}=0), which were obtained from the previous step (Table 2). As output, the models predict the value of a binary variable (p_{CT}) related to the presence of chronic lung lesions on CT images, with p_{CT}=1 for a CT score ≥ 7 (n=129) and p_{CT}=0 for a CT score < 7 (n=128) (Figure 1).

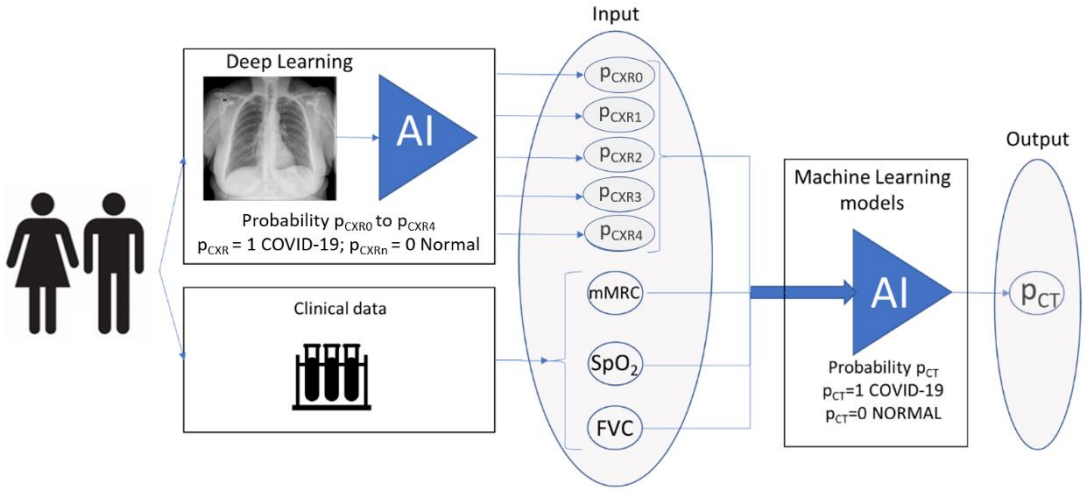


Figure 1. Machine learning-based model. Data on the modified Medical Research Council (mMRC) dyspnea scale, oximetry (SpO2), and spirometry (forced vital capacity [FVC]), and radiographic probabilities (p_{CXR0} to p_{CXR4}) with findings related to COVID-19 (p_{CXRn}=1) and normal (p_{CXRn}=0) were used as input variables, and the presence of lung lesions due to COVID-19 (p_{CT}) was used as output. AI, artificial intelligence. CT, computed tomography.

The first model was a logistic regression (LR) model with L2 regularization to prevent overfitting,[4] whereas the second model was a random forest model with 100 trees (RF-100), Gini criterion, minimum of two samples for splitting, minimum of one sample in leaves, and bootstrap.[4] The third model was a random forest model with parameters as described above, except for the limit of 10 trees and maximum depth h_max=6 (RF-10).[4] The performance of the machine-learning models was evaluated based on sensitivity, specificity, AUC, and F1-score.

Three combinations of input variables were evaluated: 1) clinical variables (mMRC, SpO_2, and FVC); 2) CXR; and 3) clinical variables (mMRC, SpO_2, FVC) and CXR.

For each model, a five-fold cross-validation strategy was adopted for the training and validation sets. The performance of the LR model was better when a combination of all variables (clinical variables and CXR) was used. The following metrics expressed in terms of mean ± standard deviation and 95% Confidence Interval (CI) were considered: sensitivity, 0.85±0.08 (95% CI [0.77, 0.94]); specificity, 0.70±0.14 (95% CI [0.55, 0.85]); F1-score, 0.79±0.06 (95% CI [0.73, 0.85]); and AUC, 0.80±0.07(95% CI [0.72, 0.87]) (Table 3).

Table 3. Predictive performance of three multivariate models using three datasets.

Groups of variables	Method	Sensitivity	Specificity	F1-score	AUC
1 SpO_2, mMRC score, and FVC	LR	0.87±0.16	0.42±0.33	0.71±0.03	0.68±0.10
	RF-10	0.88±0.15	0.37±0.32	0.71±0.03	0.66±0.08
	RF-100	0.82±0.12	0.44±0.13	0.69±0.08	0.62±0.12
2 CXR	LR	0.88±0.05	0.52±0.14	0.75±0.04	0.78±0.05
	RF-10	0.91±0.08	0.41±0.18	0.73±0.04	0.73±0.06
	RF-100	0.94±0.07	0.33±0.19	0.72±0.03	0.72±0.03
3 SpO_2, mMRC score, FVC and CRX	LR	**0.85±0.08**	**0.70±0.14**	**0.79±0.06**	**0.80±0.07**
	RF-10	0.85±0.09	0.61±0.22	0.76±0.04	0.76±0.08
	RF-100	0.89±0.06	0.49±0.17	0.75±0.04	0.76±0.07

Values are presented as the mean ± standard deviation after five-fold cross validation for each test fold. Area under the receiver operating characteristic curve (AUC); Accuracy (Acc); Chest X-Ray (CRX); Forced vital capacity (FVC); Logistic Regression (LR); modified Medical Research Council dyspnea scale (mMRC); Precision (Prec); Random forest (RF).

The LR model is represented by the following function:

$$p_{CT} = \sigma\left(\beta_1 FVC^* + \beta_2\, mMRC^* + \beta_3 S_p O_2 + \beta_4 p_{CXR0} + \beta_5 p_{CXR1} + \beta_6 p_{CXR2} + \beta_7 p_{CXR3} + \beta_8 p_{CXR4}\right)$$

$$\beta_1 = -0.3705 \;\; \beta_2 = -2.2807 \;\; \beta_3 = -0.745 \;\; \beta_4 = 1.1257$$

$$\beta_5 = 1.4960 \;\; \beta_6 = 1.0761 \;\; \beta_7 = 0.7328 \;\; \beta_8 = -0.7613$$

where p_{CT} is the probability of the presence of abnormalities on CT images, σ is the sigmoid function to restrict p_{CT} between 0 and 1, $FVC^* = \frac{FVC_{Resting}}{2FVC_{min}}$, $mMRC^* = \frac{mMRC}{4}$, and p_{CXR0} to p_{CXR4} are the probabilities that the CXR image has findings related to sequelae from COVID-19, obtained in each fold (0 to 4) during a 5-folds cross validation. Table 4 shows the estimates for the logistic regression function.

Table 4. Estimates of the logistic regression function.						
Variable	**Estimated regression coefficient (β)**	**Estimated Standard Error**	***p*-value**	**95% CI for regression coefficient (β)**		**Estimated odds ratios**
FVC^*	-0.3705	0.3210	0.248	-0.9990	0.2580	0.6904
$mMRC^*$	-2.2807	0.3020	<0.001	-2.8730	-1.6890	0.1022
S_pO_2	-0.7450	0.2320	0.001	-1.2010	-0.2890	0.4747
p_{CXR0}	1.1257	0.4150	0.007	0.3120	1.9400	3.0824
p_{CXR1}	1.4960	0.4160	<0.001	0.6810	2.3110	4.4638
p_{CXR2}	1.0761	0.3390	0.002	0.4120	1.7410	2.9332
p_{CXR3}	0.7328	0.3380	0.030	0.0710	1.3950	2.0809
p_{CXR4}	-0.7613	0.4580	0.096	-1.6590	0.1360	0.4671
Forced vital capacity (FVC); modified Medical Research Council dyspnea scale (mMRC); radiographic probabilities (Pcxr0 to Pcxr4).						

Also, we included demographic and anthropometric variables on the logistic regression prediction model, performing experiments using six different combinations of variables (age, gender, body mass index [BMI], SpO2, mMRC score, FVC and CXR). The performance of each combination is reported in the Table 5. The model performance with the inclusion of demographic or anthropometric variables did not result in significant improvement. According to our experiments, the combination of SpO2, mMRC score, FVC and CXR presented the best performance.

Table 5. Performance of the predictive model using six combinations of variables (N=257).				
Groups of variables	**Sensitivity**	**Specificity**	**F1-score**	**AUC**
1 Age, gender, and BMI	0.87±0.09	0.40±0.27	0.71±0.03	0.64±0.09
2 SpO_2, mMRC score, and FVC	0.87±0.16	0.42±0.33	0.71±0.03	0.68±0.10
3 Age, Gender, BMI, SpO2, mMRC score, and FVC	0.95±0.05	0.37±0.30	0.75±0.06	0.71±0.10
4 CXR	0.88±0.05	0.52±0.14	0.75±0.04	0.78±0.05
5 Age, Gender, BMI, SpO_2, mMRC score, FVC, and CXR	0.87±0.08	0.65±0.16	0.79±0.06	0.79±0.06
6 SpO_2, mMRC score, FVC, and CXR	0.85±0.08	0.70±0.14	0.79±0.06	0.80±0.07
Values are presented as the mean ± standard deviation after five-fold cross validation for each test fold. Area under the receiver operating characteristic curve (AUC); Body Mass Index (BMI); Chest X-Ray (CRX); Forced vital capacity (FVC); modified Medical Research Council dyspnoea scale (mMRC).				

Dataset and normalization of clinical data

A total of 257 patients with data on the mMRC dyspnea scale, oximetry, spirometry, CRX, and chest CT were selected to predict pulmonary changes. Of the 257 patients, 128 had no significant CT changes (scores < 7). A CT score of 7 was used as the cutoff value by maximizing F1 scores and AUC (Figure 2).

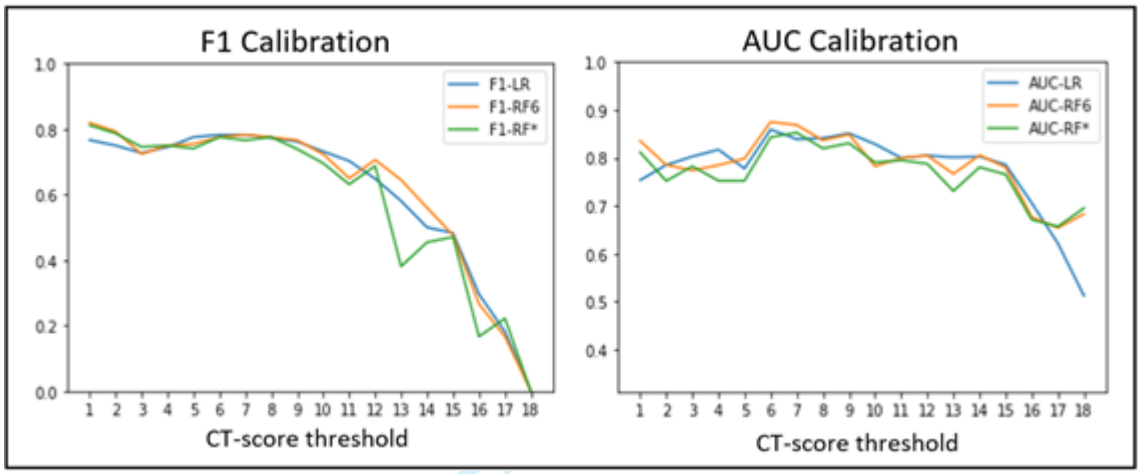
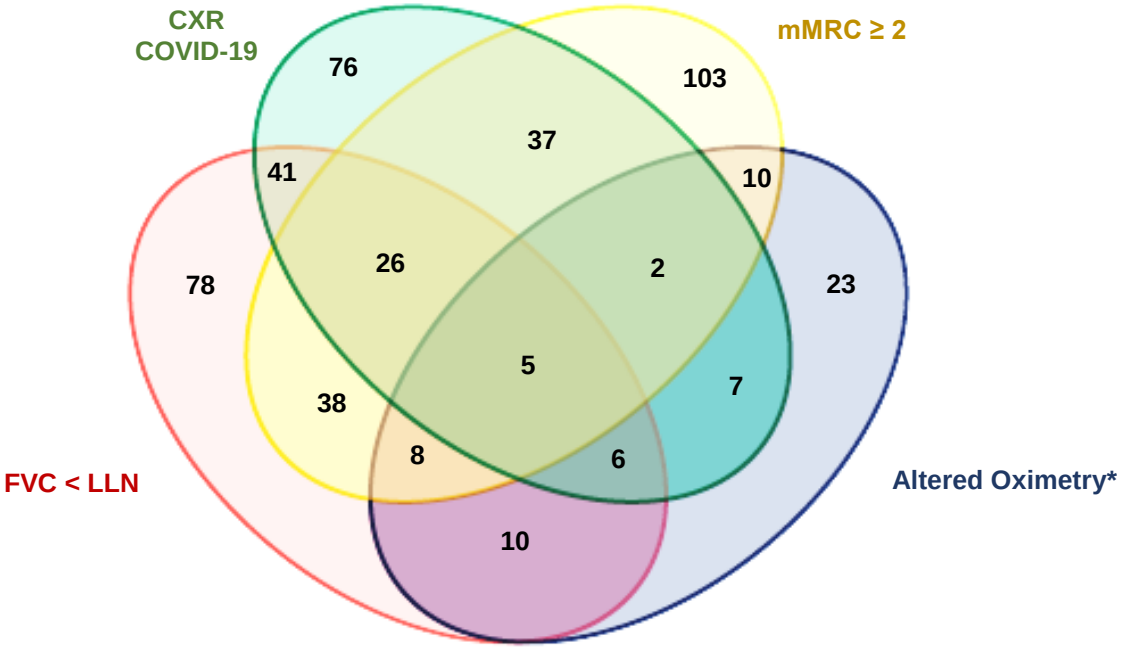
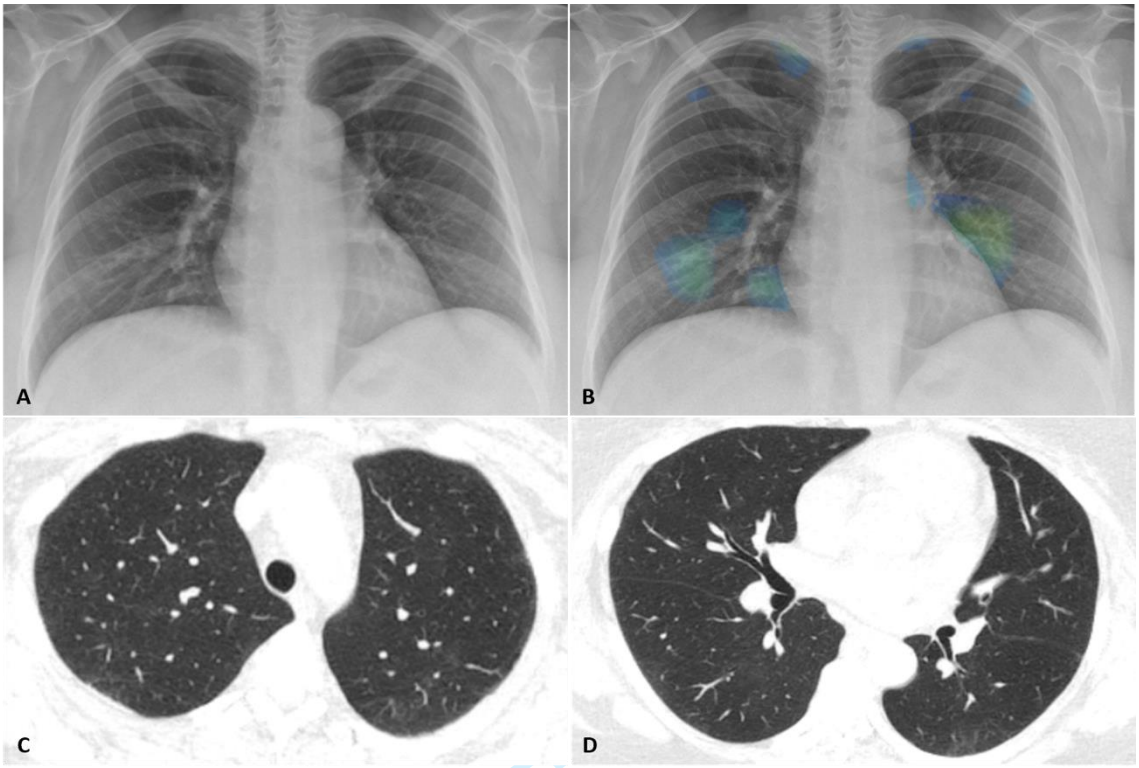


Figure 2. Computed tomography scores based on the F1-score and AUC values.

Clinical variables were normalized by dividing the mMRC values by 4 (resulting in values between 0 and 1) and the $FVC_{Resting}$ by twice the FVC_{min} (resulting in a minimum value of 0.257 and a maximum value of 0.847).



Supplemental Figure S1. Diagram showing the overlap in the changes of parameters used as pulmonary criteria to refer patients for thorax computed tomography. Values are expressed as the number of patients showing the correspondent alterations. CXR, chest X-Ray; FVC, forced vital capacity; LLN, lower limit of normal; mMRC, modified Medical Research Council dyspnea scale. *Resting SpO_2 ≤ 90% or a decrease in SpO_2 of ≥ 4% during the 1 min sit-and-stand test.



Supplemental Figure 2. Representative scan of a patient in her late 40s showing resolving ground glass abnormality after moderate COVID-19. (A) PA chest radiograph obtained 8 months after admission was considered normal by radiologists. (B) The same radiograph analyzed by the AI algorithm with heat map. Small focal abnormalities in the apical and paracardiac regions of the lungs are highlighted in green and blue. (C, D) Chest CT obtained 11 months after admission shows mild residual ground glass abnormality in the periphery of the upper lobes and left lower lobe. The patient complained of dyspnea (mMRC=3) but had normal lung function (FVC=3.81 L/91% pred) and normal oximetry (99%).

Supplemental Table S1. Demographic and clinical characteristics of the cohort of post-COVID-19 patients in this study (N=749).

Variables	Values
Age (years)	56.1 (44.4–65.1)
Male sex	399 (53.3)
BMI (kg/m^2)	30.8 (27.7–35.6) {746}
Comorbidities	
Hypertension	425 (56.7)
Smokers	285/743 (38.4)
Diabetes	261 (34.8)
COPD	55 (7.3)
Admission	
ICU	445 (59.4)
Length of ICU stay (days)	10 (6–18) {445}
IMV	304/445 (68.3)
Vital signs	
Body temperature (°C)	36.1 (35.6–36.0) {748}
Systolic blood pressure (mmHg)	124 (116–135) {743}
Diastolic blood pressure (mmHg)	77 (70–84) {743}
Heart rate (bpm)	73 (67–83) {747}
Respiratory rate (rpm)	20 (18–2) {736}
Oxygen saturation (%)	97 (95.2–98) {746}

Values are presented as median (IQR), median (IQR) {n}, n (%), or n/N (%). COPD, chronic obstructive pulmonary disease; BMI, body mass index; ICU, intensive care unit. IMV, invasive mechanical ventilation.

Supplemental Table S2. Demographic and clinical characteristics of patients with and without pulmonary involvement (N=749).

Variables	Pulmonary involvement (n=470)	No pulmonary involvement (n=279)	*p*-value
Age (years)	57.9 (45.7–65.8)	53.9 (42.5–63.7)	0.000
Male sex	228 (48.5)	171 (61.3)	0.001
BMI (kg/m^2)	31.2 (27.7–35.9) {469}	30.5 (27.6–35.2) {277}	0.111
Comorbidities			
Hypertension	287 (61.1)	138 (49.5)	0.000
Smokers	188/468 (40.2)	97/275 (35.3)	0.104
Diabetes	179 (38.1)	82 (29.4)	0.009
COPD	42 (8.9)	13 (4.7)	0.044
Admission			
ICU	317 (67.4)	128 (45.9)	0.000
Length of ICU stay (days)	11 (6–20) {317}	8 (4–14) {128}	0.000
IMV	222/317 (70)	82/128 (64.1)	0.260

Values are presented as median (IQR), median (IQR) {n}, n (%), or n/N (%). COPD, chronic obstructive pulmonary disease; BMI, body mass index; ICU, intensive care unit. IMV, invasive mechanical ventilation.

Supplemental Table S3. Demographic and clinical characteristics of COVID-19 patients with signs of pulmonary involvement (N=470).

Variables	Patients with signs of pulmonary involvement		*p*-value
	Those who underwent CT (n=348)	Those who did not undergo CT (n=122)	
Age (years)	57.8 (45.7–65.8)	58.1 (45.3–65.8)	0.490
Male sex	163 (46.8)	65 (53.3)	0.392
BMI (kg/m^2)	31.6 (28.0–36.0)	30.3 (27.0–35.9) {121}	0.041
Comorbidities			
Hypertension	215 (61.8)	72 (59)	0.469
Smokers	139/347 (40.1)	49/121 (40.5)	0.762
Diabetes	142 (40.8)	37 (30.3)	0.999
COPD	32 (9.2)	10 (8.2)	0.826
Admission			
ICU	237 (68.1)	80 (65.6)	0.999
Length of ICU stay (days)	11 (6–20) {237}	10 (4.7–19) {80}	0.913
IMV	174/237 (73.4%)	48/80 (60%)	0.034

Values are presented as median (IQR), median (IQR) {n}, n (%), or n/N (%). COPD, chronic obstructive pulmonary disease; BMI, body mass index; ICU, intensive care unit. IMV, invasive mechanical ventilation.

Supplemental Table S4. Chest computed tomography (CT) features in COVID-19 patients with CT score ≥ 7 (N=156).	
Variables	**CT changes**
CT score ≥ 7	156/328 (47.6)
Characteristics (n=156)	
Ground-glass opacities	153 (98.1)
Parenchymal bands	143 (91.7)
Reticulations	134 (85.9)
Traction bronchiectasis	92 (59)
Architectural distortion	73 (46.8)
Perilobular opacities	50 (32.1)
Bronchial wall thickening	38 (24.4)
Mosaic attenuation pattern	32 (20.5)
Consolidations	3 (1.9)
Pneumatocele	2 (1.3)
Honeycombing	-
Of the 328 patients who underwent CT scan, 47.6% had a CT score ≥ 7. Values are n/N (%) or n (%).	

Supplemental Table S5. Computed tomography changes 6 to 11 months after hospitalization due to COVID-19 (N=328).

Characteristics	Total cohort (N=328)	ICU Patients (N=222)	Ward Patients (N=106)
Ground-glass opacities	251 (76.5)	197 (86.6)	54 (51.3)
Parenchymal bands	209 (63.7)	169 (76.5)	40 (41)
Reticulations	169 (51.5)	145 (66.5)	24 (23.1)
Traction bronchiectasis	98 (29.9)	91 (44.1)	7 (7.7)
Architectural distortion	78 (23.8)	73 (35.8)	5 (6.4)
Bronchial wall thickening	89 (27.1)	60 (27.4)	29 (25.6)
Mosaic attenuation pattern	58 (17.7)	46 (20.1)	12 (11.5)
Perilobular opacities	50 (14)	47 (24.6)	3 (2.6)
Consolidation	3 (0.9)	3 (1.7)	-
Pneumatocele	2 (0.6)	2 (1.1)	-
Honeycombing	-	-	-

Values are presented as n (%).

Supplemental Table S6. Demographic and clinical characteristics of COVID-19 patients with pulmonary involvement stratified by inclusion in prediction analysis of pulmonary changes (N=328).

Variables	Patients with Pulmonary Changes		*p*-value
	Included Patients (N=257)	Excluded Patients (N=91)	
Age (years)	56.5 (45.7–64.4)	60.5 (46.9–69.9)	0.011
Male sex	113 (44)	50 (54.9)	0.068
BMI (kg/m^2)	32 (28.8–36.8)	30.6 (26.8–35.4)	0.054
Comorbidities			
Hypertension	151 (58.7)	64 (70.3)	0.060
Smokers	97/256 (37.9)	42 (46.1)	0.173
Diabetes	103 (40.1)	39 (42.9)	0.710
COPD	20 (7.8)	12 (13.2)	0.141
Admission			
ICU	179 (69.6)	58 (63.7)	0.359
Length of ICU stay (days)	12 (6–20.5) {179}	9.5 (6.2–19.7) {58}	0.209
IMV	140 (54.7)	35 (38.6)	0.010

Values are presented as median (IQR), median (IQR) {n}, n (%), or n/N (%). BMI, body mass index; COPD, chronic obstructive pulmonary disease; ICU, intensive care unit. IMV, invasive mechanical ventilation.

Supplemental References

1. Stephens K. SIIM, FISABIO, and RSNA Host Machine Learning Challenge for COVID-19 Detection and Localization. . *AXIS Imaging News* . 2021.

2. Tan M, Le Q. Efficientnet: Rethinking model scaling for convolutional neural networks. *International Conference on Machine Learning* 2019. p. 6105-14.

3. Russakovsky O, Deng J, Su H, et al. ImageNet Large Scale Visual Recognition Challenge. *International Journal of Computer Vision* 2015; **115**: 211-52.

4. Vittinghoff E, Glidden DV, Shiboski SC, et al. Regression Methods in Biostatistics: Linear, Logistic, Survival, and Repeated Measures Models. 2nd ed. New York: Springer-Verlag, 2012:1272.

TRIPOD Checklist: Prediction Model Development and Validation

Section/Topic	Item		Checklist Item	Page
Title and abstract				
Title	1	D;V	Identify the study as developing and/or validating a multivariable prediction model, the target population, and the outcome to be predicted.	1
Abstract	2	D;V	Provide a summary of objectives, study design, setting, participants, sample size, predictors, outcome, statistical analysis, results, and conclusions.	2
Introduction				
Background and objectives	3a	D;V	Explain the medical context (including whether diagnostic or prognostic) and rationale for developing or validating the multivariable prediction model, including references to existing models.	4
	3b	D;V	Specify the objectives, including whether the study describes the development or validation of the model or both.	5
Methods				
Source of data	4a	D;V	Describe the study design or source of data (e.g., randomized trial, cohort, or registry data), separately for the development and validation data sets, if applicable.	5
	4b	D;V	Specify the key study dates, including start of accrual; end of accrual; and, if applicable, end of follow-up.	5
Participants	5a	D;V	Specify key elements of the study setting (e.g., primary care, secondary care, general population) including number and location of centres.	5
	5b	D;V	Describe eligibility criteria for participants.	5
	5c	D;V	Give details of treatments received, if relevant.	-
Outcome	6a	D;V	Clearly define the outcome that is predicted by the prediction model, including how and when assessed.	7
	6b	D;V	Report any actions to blind assessment of the outcome to be predicted.	Data Supplement (Pg. 3 and 5)
Predictors	7a	D;V	Clearly define all predictors used in developing or validating the multivariable prediction model, including how and when they were measured.	6, 7
	7b	D;V	Report any actions to blind assessment of predictors for the outcome and other predictors.	Data Supplement (Pg. 3 and 5)
Sample size	8	D;V	Explain how the study size was arrived at.	5
Missing data	9	D;V	Describe how missing data were handled (e.g., complete-case analysis, single imputation, multiple imputation) with details of any imputation method.	Data Supplement, (Pg. 6)
Statistical analysis methods	10a	D	Describe how predictors were handled in the analyses.	Data Supplement (Pg. 3 and 5)
	10b	D	Specify type of model, all model-building procedures (including any predictor selection), and method for internal validation.	Data Supplement (Pg. 3 and 5)
	10c	V	For validation, describe how the predictions were calculated.	Data Supplement (Pg. 3 and 5)
	10d	D;V	Specify all measures used to assess model performance and, if relevant, to compare multiple models.	Data Supplement (Pg. 3 and 5)
	10e	V	Describe any model updating (e.g., recalibration) arising from the validation, if done.	S Data Supplement (Pg. 4)
Risk groups	11	D;V	Provide details on how risk groups were created, if done.	n.a.
Development vs. validation	12	V	For validation, identify any differences from the development data in setting, eligibility criteria, outcome, and predictors.	Data Supplement (Pg. 3 and 5)
Results				
Participants	13a	D;V	Describe the flow of participants through the study, including the number of participants with and without the outcome and, if applicable, a summary of the follow-up time. A diagram may be helpful.	8
	13b	D;V	Describe the characteristics of the participants (basic demographics, clinical features, available predictors), including the number of participants with missing data for predictors and outcome.	8
	13c	V	For validation, show a comparison with the development data of the distribution of important variables (demographics, predictors and outcome).	Data Supplement (Pg. 3 and 5)
Model development	14a	D	Specify the number of participants and outcome events in each analysis.	Data Supplement (Pg. 3 and 5)
	14b	D	If done, report the unadjusted association between each candidate predictor and outcome.	Data Supplement (Pg. 3 and 5)
Model specification	15a	D	Present the full prediction model to allow predictions for individuals (i.e., all regression coefficients, and model intercept or baseline survival at a given time point).	Data Supplement (Pg. 3 and 5)

TRIPOD Checklist: Prediction Model Development and Validation

	15b	D	Explain how to the use the prediction model.	Data Supplement (Pg. 3 and 5)
Model performance	16	D;V	Report performance measures (with CIs) for the prediction model.	Data Supplement (Pg. 5)
Model-updating	17	V	If done, report the results from any model updating (i.e., model specification, model performance).	Data Supplement (Pg.5)
Discussion				
Limitations	18	D;V	Discuss any limitations of the study (such as nonrepresentative sample, few events per predictor, missing data).	12
Interpretation	19a	V	For validation, discuss the results with reference to performance in the development data, and any other validation data.	Data Supplement (Pg.5)
	19b	D;V	Give an overall interpretation of the results, considering objectives, limitations, results from similar studies, and other relevant evidence.	10, 11, 12
Implications	20	D;V	Discuss the potential clinical use of the model and implications for future research.	10, 11, 12
Other information				
Supplementary information	21	D;V	Provide information about the availability of supplementary resources, such as study protocol, Web calculator, and data sets.	n.a.
Funding	22	D;V	Give the source of funding and the role of the funders for the present study.	14

*Items relevant only to the development of a prediction model are denoted by D, items relating solely to a validation of a prediction model are denoted by V, and items relating to both are denoted D;V. We recommend using the TRIPOD Checklist in conjunction with the TRIPOD Explanation and Elaboration document.