BMJ Open

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

Carlos Roberto Ribeiro Carvalho ,[1] Rodrigo Caruso Chate,[2] Marcio Valente Yamada Sawamura,[2] Michelle Louvaes Garcia,[1] Celina Almeida Lamas,[1] Diego Armando Cardona Cardenas,[3] Daniel Mario Lima ,[3] Paula Gobi Scudeller,[1] João Marcos Salge,[1] Cesar Higa Nomura,[2] Marco Antonio Gutierrez,[1] Members of the HCFMUSP COVID-19 Study Group

[1]Instituto do Coração—Divisão de Pneumologia, Universidade de São Paulo Hospital das Clínicas, Sao Paulo, Brazil
[2]Instituto de Radiologia, Universidade de São Paulo Hospital das Clínicas, Sao Paulo, Brazil
[3]Instituto do Coração—Divisão de Informática, Universidade de São Paulo Hospital das Clínicas, Sao Paulo, Brazil

Correspondence to
Professor Carlos Roberto Ribeiro Carvalho;
carlos.carvalho@hc.fm.usp.br

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 30 March 2020 and 31 August 2020 followed-up 6 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 CT. mMRC scale, SpO_2, 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 area under the curve, 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 patients with COVID-19.
⇒ 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.

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 >12 weeks after acute infection, characterising long COVID-19.[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-19.[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 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

INTRODUCTION

COVID-19 caused by SARS-CoV-2 emerged in December 2019 and had since spread globally.[1] This multisystemic viral disease promotes endothelial and microvascular

pulmonary sequelae, which reduce functional capacity and the health-related quality of life.[6 11 12] The National Institute for Health and Care Excellence 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 characterisation of 'fibrotic-like' lung lesions.[1 10]

WHO reported >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 UK National Institute for Health Research invested £18.5 million to fund studies on long COVID-19.[16] The lack of knowledge and medical training for treating post-COVID symptoms also represents a significant public health challenge.[14] Thus, healthcare systems will have to reorganise 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 patients with COVID-19 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 radiological 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 30 March to 31 August 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 were previously described by Busatto et al[21] and was registered at the 'Brazilian Registry of Clinical Trials' (https://ensaiosclinicos.gov.br/).

The patients were invited to participate in the study 6 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 American Thoracic Society (ATS)/European Respiratory Society (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 fivefold cross-validation strategy was adopted to train and validate the model, leading to an average area under the curve (AUC) of 0.89 (online 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. 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 categorised 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] Categorisation 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 online supplemental methods.

Machine learning model

A machine learning (ML) model based on a logistic regression (LR) with L2 regularisation to prevent overfitting[28] was adopted to detect the presence of COVID-19-related chronic lung lesions. The L1 regularisation 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
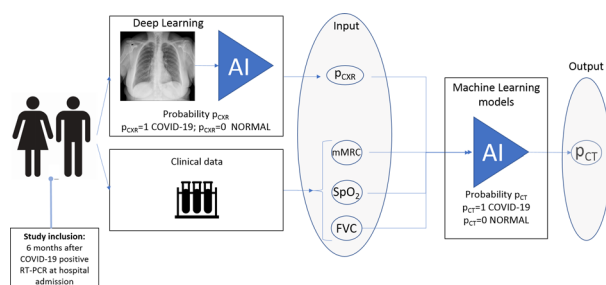

Figure 1 Logistic regression-based machine learning model to detect the presence of COVID-19-related lung lesions. The patients were invited to participate in the study 6 months after COVID-19-positive RT-PCR at hospital admission. The modified Medical Research Council (mMRC) dyspnoea scale, oximetry (SpO_2), spirometry (forced vital capacity (FVC)) and the five radiographic scores obtained during DL-based classification of chest X-ray (pCXR) were used as input data, and the presence of lung lesions due to COVID-19 was used as output data. AI, artificial intelligence.

sensitivity, specificity, AUC and F1-score after a fivefold cross-validation. (online supplemental methods)

Statistical analysis

Continuous variables are expressed as the mean and SD or median and IQR. 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 χ^2 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 curve. The performance of the ML model was determined based on sensitivity, specificity, F1-score and AUC values (online 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 3753 enrolled patients with COVID-19, 1957 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 online 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 (online 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). Online supplemental figure S1 illustrates the simultaneous presence of two or more criteria for pulmonary involvement.

The demographic and clinical characteristics of patients stratified by the presence of pulmonary involvement are described in online 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 (online 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 (online supplemental table S3).
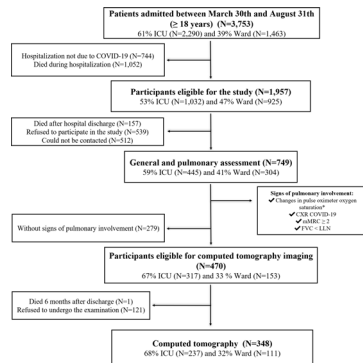
Figure 2 Flow chart of patient selection. *Rest SpO$_2$ <90% or a decrease in SpO$_2$ of at least 4% after the 1 min sit and stand test. CXR, chest X-ray; FVC, forced vital capacity; ICU, intensive care unit; LLN, lower limit of normal; mMRC, modified Medical Research Council dyspnoea scale.

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 (online 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 online 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 online 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
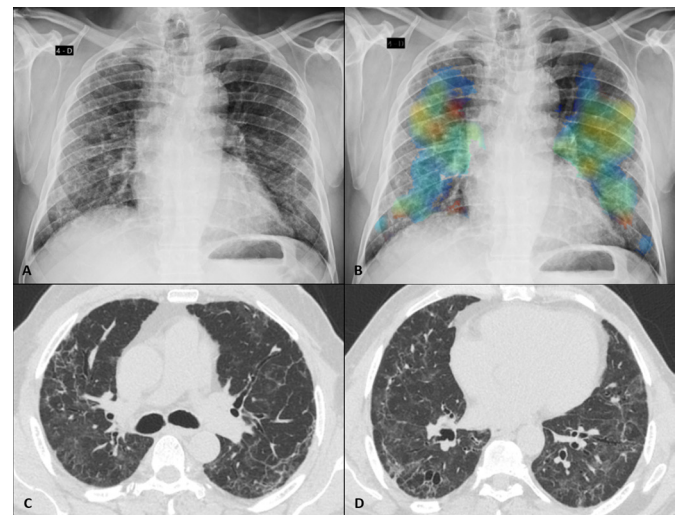

Figure 3 Fibrotic-like changes after critical COVID-19 in a patient in his early 70s. (A) Posteroanterior chest radiograph obtained 7 months after infection shows reticular opacities with a slight peripheral predominance diffusely distributed in both lungs. (B) Image from the same radiograph analysed by the artificial intelligence 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 (modified Medical Research Council dyspnoea scale=1) and altered forced vital capacity (2.34 L/60% pred), besides the normal oximetry (97%).

tests (mMRC, oximetry, spirometry, CXR and CT) and 30 showed radiographic signs not related to COVID-19 (online 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±SD and 95% CIs were observed: sensitivity, 0.85±0.08 (95% CI (0.77 to 0.94)); specificity, 0.70±0.14 (95% CI (0.55 to 0.85)); F1-score, 0.79±0.06 (95% CI (0.73 to 0.85)) and AUC, 0.80±0.07 (95% CI (0.72 to 0.87)) (table 2).

The ML predictive model is represented by the following function:

$$p_{CT} = \sigma(\beta_1 FVC^* + \beta_2 mMRC^* + \beta_3 S_p O_2 + \beta_4 p_{CXR0} + \beta_5 p_{CXR1} + \beta_6 p_{CXR2} + \beta_7 p_{CXR3} + \beta_8 p_{CXR4})$$
$$\beta_1 = -0.3705 \quad \beta_2 = -2.2807 \quad \beta_3 = -0.745 \quad \beta_4 = 1.1257$$
$$\beta_5 = 1.4960 \quad \beta_6 = 1.0761 \quad \beta_7 = 0.7328 \quad \beta_8 = -0.7613$$

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

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 (%).
*Resting SpO$_2$ ≤90% or a decrease in SpO$_2$ of ≥4% during the 1 min sit and stand test.
CXR, chest X-ray; FVC, forced vital capacity; LLN, lower limit of normal; mMRC, modified Medical Research Council dyspnoea scale.

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±SD after fivefold cross-validation for each test fold.
CXR, chest X-Ray; FVC, forced vital capacity; mMRC, Modified Medical Research Council dyspnoea scale.

in each fold (0–4) during a fivefold cross-validation (online supplemental methods).

Therefore, based on these observations, we propose in a flow chart 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 6 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 patients with COVID-19 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 hospitalised patients with COVID-19 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 patients who had COVID-19 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 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 patients with COVID-19.[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 3 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
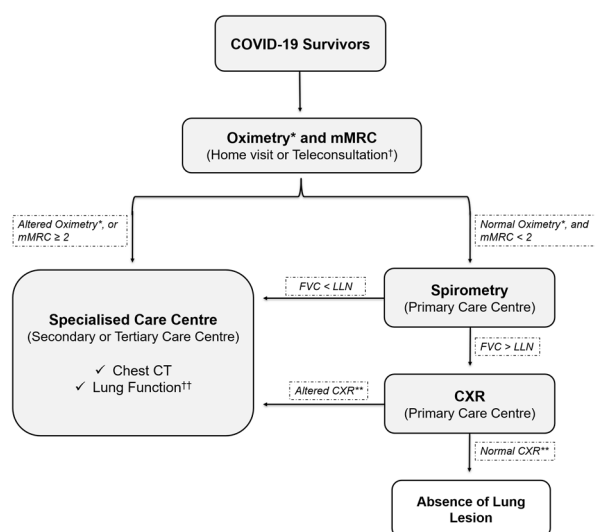

Figure 4 Flow chart for lung lesion case-finding in COVID-19 survivors. *Altered oximetry: resting SpO$_2$ ≤90% or a decrease in SpO$_2$ of ≥4% during the 1 min sit and stand test. **Altered CXR: COVID-19 findings, including bilateral linear and/or reticular opacities, especially peripheral opacities. †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.

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 patients with COVID-19 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.

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 flow chart suggested for lung lesion case-finding in COVID-19 survivors (figure 4). First, 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 healthcare 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 generalisability 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 low-income and middle-income 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.

and editing. MVYS: data curation, formal analysis, investigation, methodology, validation, visualisation, writing—original draft and writing—review and editing. MLG: conceptualisation, data curation, formal analysis, investigation, methodology, project administration, supervision, validation, visualisation, writing—original draft and writing—review and editing. CAL: data curation, formal analysis, investigation, writing—original draft and writing—review and editing. DACC: data curation, formal analysis, methodology, software, writing—original draft and writing—review and editing. DML: data curation, formal analysis, methodology, software, writing—original draft and writing—review and editing. PGS: conceptualisation, project administration, supervision, validation, visualisation and writing—review and editing. JMS: methodology, validation, visualisation and writing—review and editing. CHN: methodology, validation, visualisation and writing—review and editing. MAG: data curation, formal analysis, funding acquisition, methodology, software, supervision, validation, visualisation, writing—original draft and writing—review and 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.

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

Patient consent for publication Not applicable.

Ethics approval This study was approved by Comitê de Ética e Pesquisa HCFMUSP (Process No. 31942020.0.000.0068). Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available on reasonable request. All data relevant to the study are included in the article or uploaded as supplementary information. The raw data are not publicly available because follow-up studies will be carried out. However, data are available from the corresponding author on request and authorisation 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 (FAPESP. COVID-19 DataSharing/BR, 2021. Available: https://repositoriodatasharingfapesp.uspdigital.usp.br).

ORCID iDs
Carlos Roberto Ribeiro Carvalho http://orcid.org/0000-0002-1618-8509
Daniel Mario Lima http://orcid.org/0000-0002-7818-6103

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. doi:10.3390/ijerph18084350. [Epub ahead of print: 20 04 2021].
2. Nalbandian A, Sehgal K, Gupta A, et al. Post-acute COVID-19 syndrome. *Nat Med* 2021;27:601–15.
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;22:32.
4. Tanni SE, Fabro AT, de Albuquerque A, et al. Pulmonary fibrosis secondary to COVID-19: a narrative review. *Expert Rev Respir Med* 2021;15:791–803.
5. Macedo BRde, Garcia MVF, Garcia ML, et al. Implementation of Tele-ICU during the COVID-19 pandemic. *J Bras Pneumol* 2021;47:e20200545.
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;397:220–32.
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. [Epub ahead of print: 01 Mar 2021].
8. Fernandes PMP, Mariani AW. Life post-COVID-19: symptoms and chronic complications. *Sao Paulo Med J* 2021;139:1–2.
9. Writing Committee for the COMEBAC Study Group, Morin L, Savale L, et al. Four-Month clinical status of a cohort of patients after hospitalization for COVID-19. *JAMA* 2021;325:1525–34.
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;183:788–824.
11. Carfì A, Bernabei R, Landi F, et al. Persistent symptoms in patients after acute COVID-19. *JAMA* 2020;324:603–5.
12. Ferreira JC, Ho Y-L, Besen BAMP, et al. Protective ventilation and outcomes of critically ill patients with COVID-19: a cohort study. *Ann Intensive Care* 2021;11:92.
13. WHO. Who coronavirus disease (COVID-19) Dashboard, 2021. Available: https://covid19.who.int/ [Accessed 13 Oct 2021].
14. Wade DT. Rehabilitation after COVID-19: an evidence-based approach. *Clin Med* 2020;20:359–65.
15. Godoy CG, ECGE S, Oliveira DB. Protocol for functional assessment of adults and older adults after hospitalization for COVID-19. *Clinics* 2021;76:e3030.
16. Subbaraman N. US health agency will invest $1 billion to investigate 'long COVID'. *Nature* 2021;591:356.
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;5:7.
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. doi:10.1183/13993003.00775-2020. [Epub ahead of print: 06 08 2020].
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;34:297–307.
20. Greenhalgh T, Knight M, A'Court C, et al. Management of post-acute covid-19 in primary care. *BMJ* 2020;370:m3026.
21. Busatto GF, de Araújo AL, Duarte AJdaS, 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:e051706.
22. van den Borst B, Peters JB, Brink M. Comprehensive health assessment three months after recovery from acute COVID-19. *Clin Infect Dis* 2020 (published Online First: 2020/11/21).
23. Miller MR, Hankinson J, Brusasco V, et al. Standardisation of spirometry. *Eur Respir J* 2005;26:319–38.
24. Pereira CAdeC, Sato T, Rodrigues SC. New reference values for forced spirometry in white adults in Brazil. *J Bras Pneumol* 2007;33:397–406.
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;246:697–722.
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;30:6808–17.
28. Vittinghoff E, Glidden DV, Shiboski SC. *Regression methods in biostatistics: linear, logistic, survival, and repeated measures models*. 1272. 2nd edn. New York: Springer-Verlag, 2012.
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;2021:1–7.
30. Huang L, Yao Q, Gu X, et al. 1-year outcomes in hospital survivors with COVID-19: a longitudinal cohort study. *Lancet* 2021;398:747–58.
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;301:E438–40.
32. Han X, Fan Y, Alwalid O, et al. Six-Month follow-up chest CT findings after severe COVID-19 pneumonia. *Radiology* 2021;299:E177–86.

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.

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:E396–405.

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. doi:10.1183/13993003.03690-2020. [Epub ahead of print: 29 04 2021].

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;17:1231–7.

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;27:711–9.

38 Solomon JJ, Heyman B, JP K. CT of Postacute lung complications of COVID-19. *Radiology* 2021:211396.

39 Sonnweber T, Sahanic S, Pizzini A, *et al*. Cardiopulmonary recovery after COVID-19: an observational prospective multicentre trial. *Eur Respir J* 2021;57. doi:10.1183/13993003.03481-2020. [Epub ahead of print: 29 04 2021].

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.

41 DataSUS. MdS. SIGTAP - Sistema de Gerenciamento da Tabela de Procedimentos, Medicamentos e OPM do SUS, 2021. Available: http://sigtap.datasus.gov.br/tabela-unificada/app/sec/inicio.jsp [Accessed 03 Jan 2021].

42 Mettler FAM, Bhargavan M, Chambers M, *et al*. *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.

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
	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**
SIIM-RSNA	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
	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
InRad	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} to p_{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).
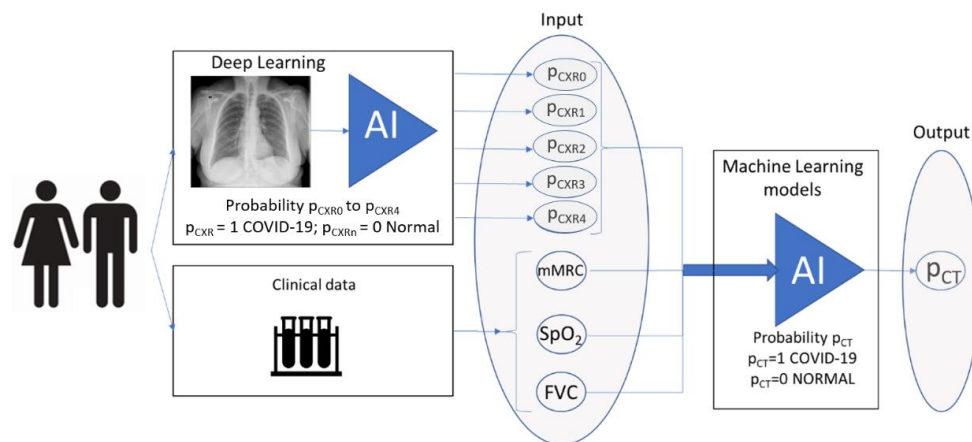
Figure 1. Machine learning-based model. Data on the modified Medical Research Council (mMRC) dyspnea scale, oximetry (SpO2), and spirometry (forced vital capacity [FVC]), and radiographic probabilities (p_{CXR0} to p_{CXR4}) with findings related to COVID-19 (p_{CXRn}=1) and normal (p_{CXRn}=0) were 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, SpO2, and FVC); 2) CXR; and 3) clinical variables (mMRC, SpO2, FVC) and CXR.

Supplemental material

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

BMJ Open

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 **SpO₂, 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 **SpO₂, 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:

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

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

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

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 **SpO2, 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, SpO2, mMRC score, FVC, and CXR**	0.87±0.08	0.65±0.16	0.79±0.06	0.79±0.06
6 **SpO2, 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).
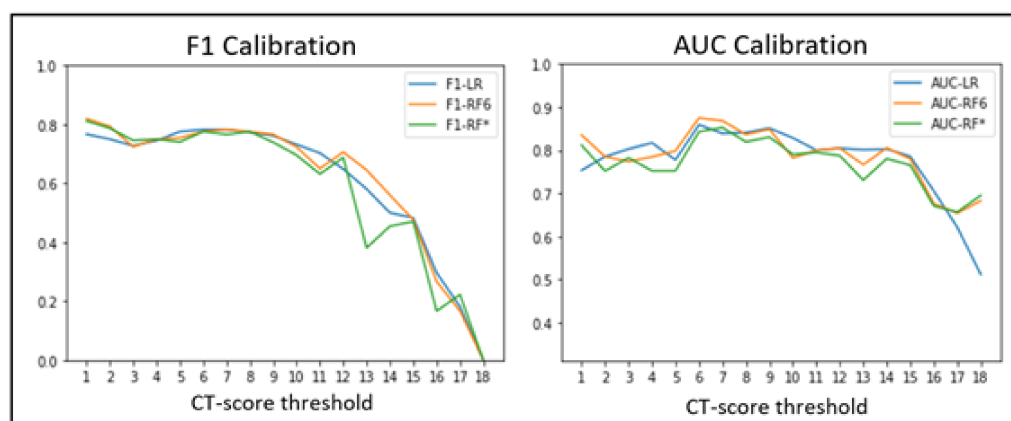
Figure 2. Computed tomography scores based on the F1-score and AUC values.

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).

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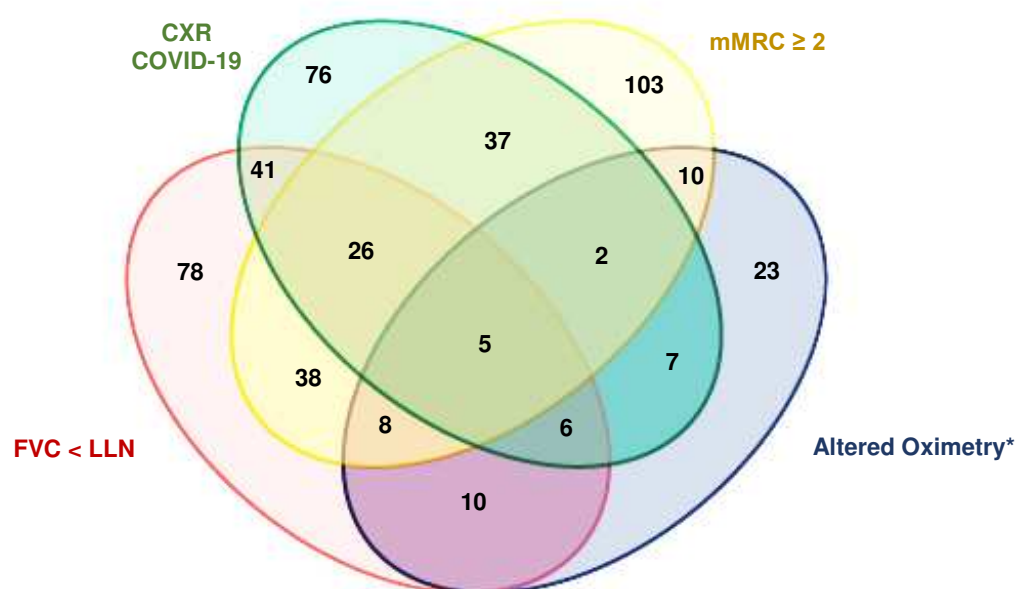
Supplemental Figure S1. Diagram showing the overlap in the changes of parameters used as pulmonary criteria to refer patients for thorax computed tomography. Values are expressed as the number of patients showing the correspondent alterations. CXR, chest X-Ray; FVC, forced vital capacity; LLN, lower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

Supplemental material

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

BMJ Open
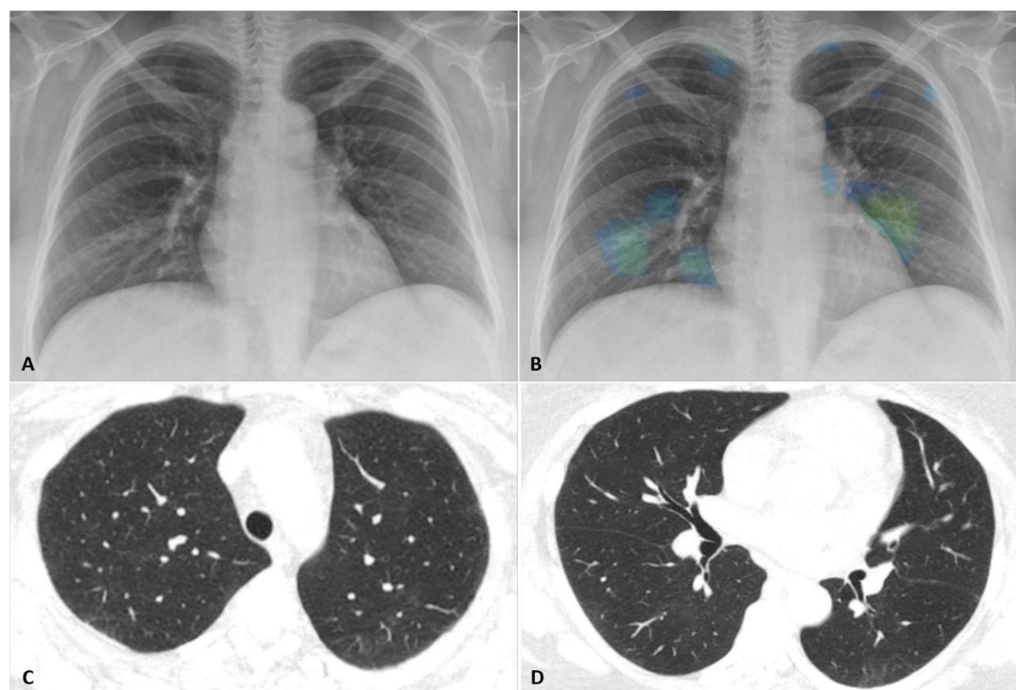
Supplemental Figure 2. Representative scan of a patient in her late 40s showing resolving ground glass abnormality after moderate COVID-19. (A) PA chest radiograph obtained 8 months after admission was considered normal by radiologists. (B) The same radiograph analyzed by the AI algorithm with heat map. 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 material

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

BMJ Open

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 material

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

BMJ Open

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 material

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance
placed on this supplemental material which has been supplied by the author(s)

BMJ Open

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.