Practical Clinical and Radiological Models to Diagnose COVID-19 on Chest CT in a French Multicentric Emergency Population

Paul Schuster  
Imadis Teleradiology

Amandine Crombé  
Imadis Teleradiology

Hubert Nivet  
Imadis Teleradiology

Alice Berger  
DeepLink Medical

Laurent Pourriol  
Imadis Teleradiology

Nicolas Favard  
Imadis Teleradiology

Alban Chazot  
Imadis Teleradiology

Florian Alonzo-Lacroix  
Imadis Teleradiology

Emile Youssof  
Imadis Teleradiology

Alexandre Ben Cheikh  
Imadis Teleradiology

Julien Balique  
Imadis Teleradiology

Basile Porta  
Imadis Teleradiology

François Petitpierre  
Imadis Teleradiology

Grégoire Bouquet  
Imadis Teleradiology

Charles Mastier  
Imadis Teleradiology

Flavie Bratan  
Imadis Teleradiology

Jean-François Bergerot  
Imadis Teleradiology

Vivien Thomson
Keywords: Logistic Models, Diagnosis, Differential, COVID-19, Teleradiology, Cohort Studies, Tomography, X-Ray Computed, Algorithms, Polymerase Chain Reaction

DOI: https://doi.org/10.21203/rs.3.rs-90375/v1

License: © This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
Abstract

Our aim was to develop practical models built with simple clinical-radiological features to facilitate COVID-19 diagnosis. To do so, 513 consecutive adult patients suspected of having COVID-19 from 15 emergency departments from 03/13/2020 to 04/14/2020 were included (244 [47.6%] with a positive RT-PCR). Chest CTs were immediately and prospectively analysed by on-call teleradiologists (OCTR) and systematically reviewed within one week by another senior teleradiologist. Each OCTR reading was concluded using a 5-point scale: normal, non-infectious, infectious non-COVID-19, indeterminate and highly suspicious of COVID-19. The senior reading reported the lesions' semiology, distribution, extent and differential diagnoses. Multivariate stepwise logistic regression (Step-LR) and classification tree (CART) models to predict a positive RT-PCR were trained on 412 patients, validated on an independent cohort of 101 patients and compared with the OCTR performances (295 and 71 with available clinical data, respectively). Regarding models elaborated on radiological variables alone, best performances were reached with the CART model (i.e., AUC=0.92 versus 0.88 for OCTR) while step-LR provided the highest AUC with clinical-radiological variables (0.93 versus 0.86 for OCTR). Hence, these two simple models, depending on the availability of clinical data, could be used by any radiologist to support their conclusion in case of COVID-19 suspicion.

Introduction

Coronavirus disease 2019 (COVID-19) is a viral disease caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). It was identified in Wuhan, China, in December 2019 and has rapidly spread worldwide. As of July 22, 2020, approximately 15 million patients have been reported in 188 countries, including 214 607 patients in France, with 30 168 deaths. Clusters of patients are regularly identified, raising the fear for a potential second wave and making health authorities and emergency departments particularly vigilant.

Real-time reverse transcription polymerase chain reaction (RT-PCR) is the gold standard for the acute diagnosis of SARS-CoV-2 in upper and lower respiratory specimens, despite possible inaccurate results (false negative and false positive). However, many hospitals do not have access to molecular testing, and the delay to get the results can be days long because of the lack of materials and technicians. Conversely, chest CT is routinely performed at most institutions and can provide a result, or at least a diagnostic orientation, in less than an hour. Previous studies have highlighted typical COVID-19 patterns consisting of peripheral, multifocal ground-glass opacities (GGO), with a sensitivity of 60-98%. Since March 2020, the French Society of Radiology (SFR) has published several educational webinars and a standardized report including a four-point scale to categorize the risk of COVID-19 on chest CT, namely: highly suspicious, compatible/indeterminate, not suspicious and normal.

In parallel, machine- and deep-learning models have been issued to facilitate and even automate the diagnosis of COVID-19 on chest CT, but their implementation in practice requires either time or mathematical and computer sciences skills or graphics processing units. Alternatively, Qin et al. have proposed a simple scoring system based on biological, clinical and radiological data with high performance (area under the ROC curve [AUC] = 0.91 in the independent validation cohort), but the added value to classical radiological assessment was not detailed. Furthermore, some information required to run the model, notably a history of exposure or the leukocyte count, may not be systematically known by radiologists.

IMADIS Teleradiology is a French company dedicated to remote interpretation of emergency CT and MRI examinations. As of March 2020, IMADIS Teleradiology had partnerships with the emergency and radiological departments of 69 hospitals covering all French regions. During the coronavirus crisis, IMADIS has been widely
involved in the diagnosis of COVID patients by remote interpretations of chest CT scans from partner centres. A standardized and dedicated workflow for patients suspected of having COVID-19 and COVID-19-dedicated webinars were thus specifically developed.

Based on our experience, the first aim of our study was to elaborate and validate practical and simple models that could be used by any radiologist in any institution, without any mathematical background, based on easily available clinical and radiological features. The second aim was to correlate the results of our models with those of the OCTR's standardized conclusions.

**Materials And Methods**

**Study design**

This prospective-retrospective observational multicentric study was approved by the French Ethics Committee for the Research in Medical Imaging (CERIM) review board (IRB CRM-2007-107) according to good clinical practices and applicable laws and regulations. The written informed consent was waived due to the nature of the analysed data, which were anonymized healthcare data. All methods were performed in accordance with the relevant guidelines and regulations. The need for written informed consent was waived because of its retrospective nature.

IMADIS Teleradiology is a French company dedicated to remote interpretation of emergency CT and MRI examinations. As of March 2020, IMADIS Teleradiology had partnerships with the emergency and radiological departments of 69 hospitals covering all French regions. The panel of IMADIS teleradiologists consisted of 109 senior radiologists with at least 5 years of emergency imaging experience (mean length of practice, 7 years) and 55 junior radiologists (i.e., residents) with 3–5 years of emergency imaging experience (mean length of practice, 4 years). Teleradiologists were on-call in groups of at least two teleradiologists per night in each of the two interpretation centres (Bordeaux and Lyon, France). All radiological reports involving COVID-19 made by junior teleradiologists were collegially validated.

Our study included all consecutive adult patients from 03/13/2020 to 04/14/2020 from 15/69 (21.7%) partner hospitals that regularly provided the RT-PCR results to IMADIS, as these patients fulfill the following inclusion criteria: need for chest CT due to suspicion of COVID-19 according to a board-certified emergency physician, available chest CT, and available RT-PCR status (Figure 1, Supplemental Data 1).

**Teleradiological Interpretation Protocol**

The IMADIS teleradiology interpretation protocol met the French recommendations for teleradiology practice. Reports and requests with clinical data for COVID-19 Chest CT image interpretation were received from the client hospitals at the IMADIS Teleradiology centres by using teleradiology software (ITIS; Deeplink Medical, Lyon, France). The images were securely transferred over a virtual private network to a local picture archiving and communication system for interpretation (Carestream Health 12, Rochester, NY). Images were immediately interpreted by OCTRs.

CT examinations were systematically reviewed within one week following each on-call session by another senior teleradiologist (n = 15; mean length of practice, 12.1 years; mean number of reviews, 34 CTs) who was not involved during the on-call duty period, blinded to the RT-PCR result and the first reader's report. All senior radiologists had a 2-hour-long e-learning session on CT-Chest findings in COVID-19, which became publicly available on April 7 (Web-based e-learning, developed by IMADIS Radiologists, Deeplink Medical, Lyon, France and RiseUp, Paris, France: https://covid19-formation.riseup.ai/).
Clinical Data

Clinical information was prospectively provided by emergency physicians and collected through the teleradiology software in a standardized COVID-19 CT request form (ITIS; Deeplink Medical, Lyon, France), as follows: age; gender; active smoking; medical history, recent anti-inflammatory drugs intake; delay since onset of symptoms (categorized as: <1 week, 1-2 weeks, ≥ 2 weeks); oxygen saturation (categorized as: ≥ 95%, 90-95% and < 90%); dyspnoea; fever (³ 38°C); cough; asthenia; headache; and ear, nose and throat symptoms.

The RT-PCR results from throat swab samples contemporary of the emergency room visit were retrospectively collected from the patients’ electronic medical records from each partner hospital.

Radiological Data

At the end of the report, the OCTR had to propose a conclusion adapted from the SFR classification, as follows: (1) normal, (2) abnormalities inconsistent with pulmonary infection; (3) abnormalities consistent with a non-COVID-19 infection; (4) indeterminate/compatible abnormalities; and (5) findings strongly suspicious of COVID-19.

The 2nd reading assessed the following radiological features: (a) underlying pulmonary disease (categorized as: emphysema, lung cancer, interstitial lung disease, pleural lesions, or bronchiectasis); (b) GGO pattern (categorized as: rounded or non-rounded GGO); (c) consolidation pattern (categorized as: rounded or non-rounded consolidations and fibrotic bands); (d) predominant pattern (categorized as: GGO or consolidation); (e) distribution pattern of lesions (categorized as: peripheral predominant, central predominant, or mixed); (f) bilateral lesions; (g) diffuse lesions (i.e., five lobes involved); (h) basal predominant lesions; (i) pleural effusion (categorized as: uni- or bilateral); (j) adenomegaly (defined as lymph node with short axis > 10 mm); (k) bronchial wall thickening (further categorized as lobar/segmental or diffuse); (l) airway secretions; (m) tree-in-bud micronodules, and (n) pulmonary embolism. Images for each radiological feature can be found in Supplemental Data 2.

Statistical Analysis

Statistical analyses were performed using R (version 3.5.3, R Foundation for Statistical Computing, Vienna, Austria). A p-value of less than 0.05 was deemed significant.

Univariate associations between clinical-radiological categorical variables and RT-PCR status were evaluated with Pearson $\chi^2$ or Fisher exact tests, except for age, which was compared between the two groups with Student’s t-test. Classification and regression models are negatively affected by high correlations between explanatory variables; hence, correlations between variables were evaluated with Spearman’s test. For each significantly correlated pair of dummy variables extracted from the same initial multilevel categorical variable, the variable with the lowest p-value at univariate analysis was selected for the multivariable modelling.

Next, the study population was randomly partitioned into a training cohort (n = 412/513, ≈ 80%) and a validation cohort (n = 101/513, ≈ 20%), with a same prevalence of RT-PCR positivity. We focused on two simple classifiers that do not require any computing interface to extract the probability for a positive RT-PCR, namely: classification and regression tree (CART, “rpart” package) and stepwise backward-forward binary logistic regression (Step-LR – minimizing the Akaike information criterion, “MASS” package). The models were built on the training cohort based on (i) either all dichotomized radiological variables or (ii) all dichotomized clinical + radiological variables - with a p-value < 0.05 at univariable analysis. The CART algorithm has a hyperparameter (i.e., a parameter that is set before the model building, while classical parameters are derived during the model building), named ‘complexity’, which
controls the size of the tree and was selected following a cross-validation step in the training cohort as minimizing the classification error rate. Next, the tree was pruned following this optimal complexity hyperparameter. The minimal number of observations in the terminal node and the splitting criteria were set to 3 and the Gini index, respectively.

Models were evaluated and compared between themselves and the prospective conclusions made by the OCTRs on the validation cohort, according to AUC. Accuracy (number of correctly classified patients divided by the total number of patients), sensitivity, specificity, negative predictive value (NPV) and positive predictive value (PPV) were estimated after dichotomizing predicted probabilities per a cut-off of 0.5. All results were given with a 95% confidence interval (95%CI). AUCs were compared using the pairwise Delong test (‘pROC’ package).

Finally, we applied a decision curve analysis (DCA) to assess the clinical usefulness of the final models in the validation cohort. DCA consists of plotting the net benefit of applying the model for clinically reasonable risk thresholds compared with two alternative strategies: (i) to treat all patients as affected by COVID-19 or (ii) to treat none of the patients. Herein, the net benefit of our models refers to the correct identification of patients with a positive or a negative RT-PCR, and the risk threshold can be seen as the harm-to-benefit ratio or the risk at which patients are indifferent about COVID-19. Hence, a low risk threshold would correspond to patients who are particularly worried about the disease.

**Results**

Table 1 summarizes the descriptive features of the study population. Overall, 513 patients were included, with a median age of 68.4 years (range: 18-100) and 241/513 women (47%). The prevalence of RT-PCR positivity was 244/513 (47.6%).

**Univariate analysis (Table 2)**

The following clinical variables were associated with RT-PCR positivity (RT-PCR+): delay since onset of symptoms ≥ 1 week, oxygen saturation < 95%, oxygen saturation < 90%, presence of fever, cough, asthenia and myalgia (p = 0.04, 0.03, 0.005, <0.001, 0.02, 0.001 and 0.008, respectively). The following radiological variables were positively associated with RT-PCR+: presence of GGO, non-rounded GGO, rounded GGO, presence of consolidation, non-rounded consolidation, fibrotic bands, intralobular reticulations, fibrosis, GGO predominant pattern, peripheral predominant location, bilateral lesions, diffuse lesions, basal predominant lesions, and low, moderate and high extent of abnormalities (all p-values <0.001).

The following radiological variables were negatively correlated with RT-PCR+: consolidation predominant pattern, central predominant location, mixed predominant location, airway secretion, bronchial wall thickening, either lobar/segmental or diffuse, and tree-in-bud micronodules (p = 0.02, 0.001, 0.002, <0.001, <0.001, <0.001, <0.001, <0.001).

**Multivariate models**

The correlation matrix of the relevant dichotomized variables is shown in Figure 2. Analysis of the correlations between similar explanatory variables enabled the selection of ‘presence of GGO’ (over ‘rounded GGO’ and ‘non-rounded GGO’), ‘peripheral predominant location’ (over ‘central predominant location’ and ‘mixed predominant location’), ‘fibrotic band consolidation’ (over ‘non-rounded consolidation’ and ‘presence of consolidation’), ‘moderate
to severe extension' (over 'low to severe extension' and 'severe extension'), and 'bronchial wall thickening' (over 'diffuse bronchial wall thickening' and 'focal/segmental bronchial wall thickening'). Thus, the total numbers of variables ultimately entered in the multivariate radiological and clinical-radiological models were set to 13 and 19 dichotomized variables, respectively.

Figure 3 shows the final decision trees relying on radiological and clinical-radiological variables. Table 3 shows the tables enabling the calculation of the probability for RT-PCR+ according to the final Step-LR models relying on radiological and clinical-radiological variables.

**Performances of the models and OCTRs’ conclusions**

To evaluate the performances of the models, trained models were tested on the external validation cohort. Table 4 shows the results (results in the training cohort are given in Supplemental Data 3). Regarding radiological models, the highest performances were reached with the CART model ($AUC_{\text{validation}} = 0.91$, 95%CI = [0.86 – 0.98] and $\text{accuracy}_{\text{validation}} = 0.86$, 95%CI = [0.78 – 0.92], versus $AUC_{\text{validation}} = 0.90$, 95%CI = [0.84 – 0.96] and $\text{accuracy}_{\text{validation}} = 0.83$, 95%CI = [0.74 – 0.90] in Step-LR model). Regarding clinical-radiological models, the highest performances were reached with the Step-LR model ($AUC_{\text{validation}} = 0.93$, 95%CI = [0.87 – 0.98] and $\text{accuracy}_{\text{validation}} = 0.86$, 95%CI = [0.76 – 0.93], versus $AUC_{\text{validation}} = 0.90$, 95%CI = [0.83 – 0.97] and $\text{accuracy}_{\text{validation}} = 0.83$, 95%CI = [0.72 – 0.91] in CART model). In the whole study population and in the validation subcohort, the AUCs of the OCTRs’ conclusions (sub-categorized as (1-2-3) or (4) or (5)) were 0.89 (95%CI = [0.86 – 0.92]) and 0.88 (95%CI = [0.72 – 0.88]), respectively. The ROC curves in the validation cohort are displayed on Figure 4. None of the $AUC_{\text{validations}}$ were significantly different (lowest p-value = 0.07 regarding the comparisons between the Step-LR clinical-radiological model and the OCTRs’ conclusions).

In the subgroup of patients with an indeterminate conclusion (4), the probabilities for RT-PCR+ with the clinical-radiological Step-LR model were significantly higher for patients with RT-PCR+ than for patients with RT-PCR- (0.63 ± 0.28 versus 0.39 ± 0.27, respectively, $p = 0.004$). Figure 5 illustrates the potential application of this model for patients with an indeterminate/compatible conclusion. An excel macro is provided in Supplementary Data 4 so that the interested reader can test the Step-LR models.

Regarding the outliers of the model with the highest AUC (i.e., clinical-radiological step-LR), 3 out of the 32 patients (9.4%) with a negative RT-PCR in the validation cohort were classified positively by the model. Of these 3 patients, 2 were highly suspicious of COVID-19 and 1 was indeterminate according to the OCTRs. Conversely, 7 out of the 39 patients (17.9%) with a positive RT-PCR in the validation cohort were classified negatively by the model. Of these 7 patients, 4 had conclusions of (1), (2) or (3) according to the OCTRs.

**Clinical usefulness of the final models in the validation cohort (Fig. 6)**

The DCAs showed that the OCTRs’ conclusion and the final models added more benefit than the ‘treat all approach’ above a risk threshold of approximately 0.05. The two final models added more net benefit than the ‘treat all’ strategy and the OCTRs’ conclusion for threshold probability above 0.43.

**Discussion**

In this study, we developed practical and ready-to-use models to predict the RT-PCR status from categorical clinical and radiological variables that are routinely available or assessable by radiologists without either expertise in
thoracic imaging or computer science or additional blood samples. We purposely elaborated parsimonious models through a cautious variable selection to facilitate their applications in practice. Our best models displayed high AUCs and accuracy on the external validation cohort and could be helpful to radiologists, notably to weight indeterminate conclusions.

Our series represent a real-life multi-centric emergency population during the COVID-19 pandemic with well-balanced proportions of negative and positive RT-PCR results, making it appropriate for the development of predictive models. The clinical and radiological variables correlating with the RT-PCR status in the univariate analysis were consistent with the literature, namely, fever, asthenia, oxygen saturation, GGO, non-rounded consolidation, fibrotic bands, and intralobular reticulations with bilateral, diffuse, basal predominant and peripheral distributions.

The aim of our models was to provide rapid assistance to non-specialized radiologists who may need support to conclude or modulate his report confidently without delaying patient management. We purposely chose classification algorithms that are easily explainable and do not require additional computing time, i.e., a binary logistic regression and a classification tree. These two algorithms are often used for benchmarking purposes in machine learning before using more complex models. In preliminary exploratory analyses, we actually tested other algorithms such as random forests or categorical boosted trees, which indeed showed slightly higher AUCs but could not be visually represented and explained to radiologists. Interestingly, similar performances of our models in the training and validation cohorts highlight the lack of overfitting and their good generalizability to new patients.

Improving COVID-19 diagnosis with deep learning frameworks has already been attempted with good results, but without remarkable added value compared with the diagnostic performances of the OCTRs or our models as calculated in the independent validation set. Indeed, the AUCs mostly ranged between 0.90 – 0.95. Other studies proposed practical scores such as the PSC-19, which relies on four variables (history of exposure, leukocyte count, peripheral lesion and crazy paving patterns)12. However, this score requires a blood sample and could be of limited use when investigating a patient in a new COVID-19 cluster of patients without proven exposure. Chen et al. also combined an explainable machine learning algorithm (i.e., penalized logistic regression) and bio-clinical-radiological variables10. They built three models: bio-clinical alone, radiological alone and bio-clinical-radiological models. Surprisingly, the results of the final models in the validation cohort showed the highest performances when radiological variables were not taken into account (AUCs = 0.97, 0.81 and 0.94, respectively). Ridge and/or LASSO penalized logistic regressions have also been investigated in our preliminary data exploration but did not show added value to Step-LR. While deep learning studies trained their models in large cohorts of hundreds of patients, it should be noted that these two practical studies did not exceed one hundred patients, questioning their validity. In addition, the performances of standardized radiological conclusions were missing in all these studies involving artificial intelligence.

Our results also highlight the very good performances of radiologists in daily practice, which can be explained by the considerable increase in knowledge of the COVID-19 radiological presentations since the first peer-reviewed studies in January-February 2020, attained through open-source educational publications, webinars and recommendations issued by national and international radiological societies, which were immediately relayed to the teleradiologists working at IMADIS. Additionally, OCTRs could ask for their colleagues’ advice when facing a complicated case, as they were never alone in one of our two emergency interpretation centres during on-call duty. Therefore, though our predictive models showed comparable performances with other machine- and deep-learning models, and slightly higher performances than OCTRs’ conclusions, the gain was not striking, leading to no significant difference according to the Delong test.
To illustrate eventual clinical applications of the final models, we used DCAs, which are a popular alternative to cost-effectiveness studies. Regarding the two settings (radiological data alone and clinical-radiological data), we found similar shapes of the DCAs for the OCTRs’ conclusion and the final models. Interestingly, at worst, the net benefits of the models were equivalent to those of OCTRs (for the intermediate risk threshold) and to the ‘treat all’ strategy (for the very low risk threshold). However, the machine learning models would be complementary to the OCTRs’ conclusion and would improve the net benefit for patients from intermediate- to high-risk thresholds.

The measurements of performance in our models should also be considered with the disease prevalence during our period of inclusion (≈ 48%) and the nature of the study population. Indeed, although PPV and NPV are useful to rapidly sort patients with COVID-19 suspicion, they depend on this prevalence, which fluctuates depending on the country and public health measures.

Our study has certain limitations. First, the RT-PCR is an imperfect gold standard, which may explain why no study has ever reached perfect performance. Indeed, the analysis of the outliers of the best model in the validation cohort with fully available clinical-radiological variables (n = 10, 3 false positives and 7 false negatives) demonstrated that 6 of these patients either could have received their chest CT before the disease manifestation or were false negative by RT-PCR. This finding stresses the risk of false negatives with chest CT due to the delay between the beginning of clinical symptoms and the appearance of the COVID-19 semiology on chest CT. Second, approximately 25-33% of clinical data were missing because they were collected in real-life emergency situations. Moreover, some radiological features were not evaluated because they were published after the beginning of the COVID-19 IMADIS workflow (for instance: multifocality, thickened vessels). Finally, although the OCTRs’ conclusions were prospectively collected, the radiological features that were necessary to elaborate the predictive models were not collected in emergency situations, which could have led to the overestimation of the associations of these features with the RT-PCR status.

To conclude, we presented the largest French multicentric emergency cohort including prospective standardized reports following national recommendations. Our findings illustrate the high diagnostic performances of the OCTRs working in a teleradiological structure entirely dedicated to emergency imaging, which promoted continuous training. This setting enables us to propose free and practical models built on easily available data, which can be used by general radiologists to confidently conclude their reports in daily practice.

Declarations

Acknowledgements

We would like to sincerely thank:

- all our partner centers and their collaborators involved in this project (CH AGEN, CH AVIGNON, CH BRIANCON, CH CARCASONNE, CH CHALON, CH CHAMBERY, CH COMPIEGNE, CH LE CREUSOT, CH LOURDES, CH MACON, CH NOYON, CH PARAY-LE-MONIAL, CH ROANNE, CH ST JULIEN-EN-GENEVOIS, CH ST JOSEPH ST LUC) and especially Drs John REFAIT, Sylvie ASSADOURIAN, Julie DONETTE, Soraya ZAID, Michèle CERUTI, Bernard BRU, Judith KARSENTY, Sandrine MERCIER, Quentin METTE, Amaud VERMEERE-MERLEN, Pascal CAPDEPON, Mahmoud KAAKI, Xavier COURTOIS, and Mrs Joëlle CAUCHETIEZ.

- the support teams at Imadis Teleradiology: Audrey PARAS, Lucie VAUTHIER, Anne LE GUELLEC and Tiffany DESMURS.

- DeepLink Medical Team: Hugues LAJOIE, Marine CLOUX and Marine MOLLARD.
Author contributions statement

All authors reviewed the manuscript and approved it.

PS, HN, NB participated in conception of the work, creation of database, interpretation of images, manuscript writing

ACr participated in creation of database, interpretation of images, data analysis, development of algorithms, manuscript writing

AB participated in creation of database

LP, NF, ACh, FAL, EY, ABC, JB, BP, FP, GB, CM, FB, JFB, VT Participated in interpretation of images

GG participated in conception of the work, manuscript writing, supervision of work

Competing interests

This study did not receive any funding.

LP, NF, ACh, EY, ABC, JB, BP, CM, JFB, VT, and GG have shares in DeepLink Medical.

Data Availability

The datasets analyzed during the current study will be publicly available (Being submitted to Springer Nature Database). Details regarding the statistical analysis can be made available from the corresponding author on reasonable request.

Abbreviations

95%CI: 95% confidence interval

AUC: area under the receiver operating characteristic curves

CART: classification and regression tree

COVID-19: coronavirus disease 2019

CT: computed tomography

DCA: decision curve analysis

GGO: ground-glass opacities

NPV: negative predictive value

OCTR: on-call teleradiologists
PPV: positive predictive value

RT-PCR: real-time reverse transcription polymerase chain reaction

SARS-CoV-2: severe acute respiratory syndrome coronavirus 2

SFR: French society of radiology

Step-LR: stepwise binary logistic regression

References

1. Zhu, N. et al. A novel coronavirus from patients with pneumonia in China, 2019. *N. Engl. J. Med.* (2020) doi:10.1056/NEJMoa2001017.

2. Hopkins., J. CSSE Coronavirus COVID-19 Global Cases (dashboard). https://gisanddata.maps.arcgis.com/apps/opsdashboard/index.html#/bda7594740fd40299423467b48e9ecf6 (2020).

3. Ai, T. et al. Correlation of Chest CT and RT-PCR Testing for Coronavirus Disease 2019 (COVID-19) in China: A Report of 1014 Cases. *Radiology* (2020) doi:10.1148/radiol.2020200642.

4. Fang, Y. et al. Sensitivity of Chest CT for COVID-19: Comparison to RT-PCR. *Radiology* (2020) doi:10.1148/radiol.2020200432.

5. Yang, S. et al. Clinical and CT features of early stage patients with COVID-19: A retrospective analysis of imported cases in Shanghai, China. *European Respiratory Journal* (2020) doi:10.1183/13993003.00407-2020.

6. Kanne, J. P., Little, B. P., Chung, J. H., Elicker, B. M. & Ketai, L. H. Essentials for Radiologists on COVID-19: An Update-Radiology Scientific Expert Panel. *Radiology* (2020) doi:10.1148/radiol.2020200527.

7. Prevention., C. for D. C. and. Interim Guidelines for Collecting, Handling, and Testing Clinical Specimens from Persons Under Investigation (PUIs) for Coronavirus Disease 2019 (COVID-19). https://www.cdc.gov/coronavirus/2019-ncov/lab/guidelines-clinical-specimens.html (2020).

8. Fang, X., Li, X., Bian, Y., Ji, X. & Lu, J. Radiomics nomogram for the prediction of 2019 novel coronavirus pneumonia caused by SARS-CoV-2. *Eur. Radiol.* (2020) doi:10.1007/s00330-020-07032-z.

9. Bai, H. X. et al. AI Augmentation of Radiologist Performance in Distinguishing COVID-19 from Pneumonia of Other Etiology on Chest CT. *Radiology* (2020) doi:10.1148/radiol.2020201491.

10. Chen, X. et al. A diagnostic model for coronavirus disease 2019 (COVID-19) based on radiological semantic and clinical features: a multi-center study. *Eur. Radiol.* (2020) doi:10.1007/s00330-020-06829-2.

11. Ni, Q. et al. A deep learning approach to characterize 2019 coronavirus disease (COVID-19) pneumonia in chest CT images. *Eur. Radiol.* (2020) doi:10.1007/s00330-020-07044-9.

12. Li, L. et al. Using Artificial Intelligence to Detect COVID-19 and Community-acquired Pneumonia Based on Pulmonary CT: Evaluation of the Diagnostic Accuracy. *Radiology* (2020) doi:10.1148/radiol.2020200905.

13. Pu, J. et al. Any unique image biomarkers associated with COVID-19? *Eur. Radiol.* (2020) doi:10.1007/s00330-020-06956-w.

14. Qin, L. et al. A predictive model and scoring system combining clinical and CT characteristics for the diagnosis of COVID-19. *Eur. Radiol.* (2020) doi:10.1007/s00330-020-07022-1.
15. Site, S. F. de R. W. Qualité et sécurité des actes de téléimagerie – Guide de bonnes pratiques. http://www.sfrnet.org/sfr/professionnels/2-infos-professionnelles/05-teleradiologie/index.phtml (2020).

16. Vickers, A. J. & Elkin, E. B. Decision curve analysis: A novel method for evaluating prediction models. *Med. Decis. Mak.* (2006) doi:10.1177/0272989X06295361.

17. Pauker, S. G. & Kassirer, J. P. Therapeutic Decision Making: A Cost-Benefit Analysis. *N. Engl. J. Med.* (1975) doi:10.1056/NEJM197507312930505.

18. Vickers, A. J., van Calster, B. & Steyerberg, E. W. A simple, step-by-step guide to interpreting decision curve analysis. *Diagnostic Progn. Res.* (2019) doi:10.1186/s41512-019-0064-7.

19. Wang, D. *et al.* Clinical Characteristics of 138 Hospitalized Patients with 2019 Novel Coronavirus-Infected Pneumonia in Wuhan, China. *JAMA - J. Am. Med. Assoc.* (2020) doi:10.1001/jama.2020.1585.

20. Huang, Y., Cheng, W., Zhao, N., Qu, H. & Tian, J. CT screening for early diagnosis of SARS-CoV-2 infection. *The Lancet. Infectious diseases* (2020) doi:10.1016/S1473-3099(20)30241-3.

21. Huang, C. *et al.* Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* (2020) doi:10.1016/S0140-6736(20)30183-5.

22. Eng, J. & Bluemke, D. A. Imaging Publications in the COVID-19 Pandemic: Applying New Research Results to Clinical Practice. *Radiology* (2020) doi:10.1148/radiol.2020201724.

**Tables**

**Table 1.** Characteristics of the study population.
### Characteristics

| Characteristics                                           | No. of patients |
|------------------------------------------------------------|----------------|
| **Age**                                                   |                |
| Mean                                                       | 65.6 ± 18.8    |
| Median (range)                                             | 68.4 (18.5 - 100.1) |
| **Gender**                                                |                |
| Female                                                     | 241/513 (47%)  |
| Male                                                       | 272/513 (53%)  |
| **Active smoking**                                         |                |
| No                                                         | 359/412 (87.1%)|
| Yes                                                        | 53/412 (12.9%) |
| **Anti-inflammatory drugs intake**                         |                |
| No                                                         | 395/418 (94.5%)|
| Yes                                                        | 23/418 (5.5%)  |
| **Significant medical history**                            |                |
| No                                                         | 151/478 (31.6%)|
| Yes                                                        | 327/478 (68.4%)|
| **Pre-existing lung chronic disease on chest CT**          |                |
| No                                                         | 414/513 (80.7%)|
| Yes                                                        | 99/513 (19.3%) |

**NOTE** – Data refer to the number of patients with percentage in parentheses, except for age.

**Table 2.** Correlation between clinical and radiological features and RT-PCR status or corresponding diagnostic accuracy measurements.
| Variables                  | RT-PCR-§ | RT-PCR+§ | p value |
|---------------------------|----------|----------|---------|
| Clinical features         |          |          |         |
| Time since the onset of symptoms |          |          |         |
| < 1 week                  | 148/240 (61.7%) | 115/223 (51.6%) | 0.02*   |
| 1-2 weeks                 | 80/240 (33.3%) | 102/223 (45.7%) | 0.04*   |
| > 2 weeks                 | 12/240 (5%) | 6/223 (2.7%) | 0.3     |
| Dyspnea                   |          |          |         |
| No                        | 56/269 (20.8%) | 38/244 (15.6%) | 0.16    |
| Yes                       | 213/269 (79.2%) | 206/244 (84.4%) |         |
| Oxygen saturation         |          |          |         |
| ≥ 95%                     | 77/195 (39.5%) | 54/191 (28.3%) | 0.006*  |
| 90 - 95%                  | 82/195 (42.1%) | 77/191 (40.3%) | 0.03*   |
| < 90%                     | 36/195 (18.5%) | 60/191 (31.4%) | 0.005** |
| Cough                     |          |          |         |
| No                        | 85/269 (31.6%) | 53/244 (21.7%) | 0.02*   |
| Yes                       | 184/269 (68.4%) | 191/244 (78.3%) |         |
| Fever (≥ 38°C)            |          |          |         |
| No                        | 109/269 (40.5%) | 34/244 (13.9%) | < 0.001*** |
| Yes                       | 160/269 (59.5%) | 210/244 (86.1%) |         |
| Asthenia                  |          |          |         |
| No                        | 125/269 (46.5%) | 78/244 (32%) | 0.001** |
| Yes                       | 144/269 (53.5%) | 166/244 (68%) |         |
| Myalgia                   |          |          |         |
| No                        | 211/269 (78.4%) | 165/244 (67.6%) | 0.008*  |
| Yes                       | 58/269 (21.6%) | 79/244 (32.4%) |         |
| Headache                  |          |          |         |
| No                        | 228/269 (84.8%) | 200/244 (82%) | 0.5     |
| Yes                       | 41/269 (15.2%) | 44/244 (18%) |         |
| Ear, nose throat symptoms |          |          |         |
| No                        | 247/269 (91.8%) | 228/244 (93.4%) | 0.6     |
| Yes                       | 22/269 (8.2%) | 16/244 (6.6%) |         |
| Radiological features     |          |          |         |
| Presence of GGO       |                |                | p-value |
|-----------------------|----------------|----------------|---------|
| No                    | 180/269 (66.9%)| 15/244 (6.1%)  | < 0.001*** |
| Yes                   | 89/269 (33.1%) | 229/244 (93.9%)|      |

| Non-rounded GGO       |                |                | p-value |
|-----------------------|----------------|----------------|---------|
| Absent                | 197/269 (73.2%)| 31/244 (12.7%) | < 0.001*** |
| Present               | 72/269 (26.8%) | 213/244 (87.3%)|      |

| Rounded GGO           |                |                | p-value |
|-----------------------|----------------|----------------|---------|
| Absent                | 238/269 (88.5%)| 162/244 (66.4%)| < 0.001*** |
| Present               | 31/269 (11.5%) | 82/244 (33.6%) |      |

| Presence of consolidation |                |                | p-value |
|---------------------------|----------------|----------------|---------|
| No                        | 196/269 (72.9%)| 92/244 (37.7%) | < 0.001*** |
| Yes                       | 73/269 (27.1%) | 152/244 (62.3%)|      |

| Non-rounded consolidation |                |                | p-value |
|----------------------------|----------------|----------------|---------|
| No                         | 217/269 (80.7%)| 172/244 (70.5%)| 0.0097* |
| Yes                        | 52/269 (19.3%) | 72/244 (29.5%) |      |

| Rounded consolidation     |                |                | p-value |
|----------------------------|----------------|----------------|---------|
| No                         | 246/269 (91.4%)| 227/244 (93%)  | 0.6     |
| Yes                        | 23/269 (8.6%)  | 17/244 (7%)    |        |

| Sub-pleural band          |                |                | p-value |
|----------------------------|----------------|----------------|---------|
| No                         | 253/269 (94.1%)| 142/244 (58.2%)| < 0.001*** |
| Yes                        | 16/269 (5.9%)  | 102/244 (41.8%)|      |

| Predominant pattern       |                |                | p-value |
|----------------------------|----------------|----------------|---------|
| None                      | 150/269 (55.8%)| 11/244 (4.5%)  | < 0.001*** |
| Consolidation             | 57/269 (21.2%) | 32/244 (13.1%) |      |
| GGO                       | 62/269 (23%)   | 201/244 (82.4%)|      |

| Distribution pattern of lesions |                |                | p-value |
|-------------------------------|----------------|----------------|---------|
| None                          | 133/269 (49.4%)| 9/244 (3.7%)   | < 0.001*** |
| Peripheral predominant        | 59/269 (21.9%) | 154/244 (63.1%)|      |
| Central predominant           | 28/269 (10.4%) | 7/244 (2.9%)   |      |
| Mixed                         | 49/269 (18.2%) | 74/244 (30.3%) |      |

| Bilateral lesions            |                |                | p-value |
|-------------------------------|----------------|----------------|---------|
| No                            | 170/269 (63.2%)| 28/244 (11.5%) | < 0.001*** |
| Yes                           |                |                |      |
|                                    | Observed 1 | Observed 2 | p-value |
|------------------------------------|------------|------------|---------|
| Yes                                | 99/269 (36.8%) | 216/244 (88.5%) | < 0.001*** |
| **Diffuse lesions**                |            |            |         |
| No                                 | 231/269 (85.9%) | 71/244 (29.1%) |         |
| Yes                                | 38/269 (14.1%) | 173/244 (70.9%) |         |
| **Basal predominant lesions**      |            |            |         |
| No                                 | 209/269 (77.7%) | 145/244 (59.4%) | < 0.001*** |
| Yes                                | 60/269 (22.3%) | 99/244 (40.6%) |         |
| **Extent of lesions**              |            |            |         |
| None                               | 134/269 (49.8%) | 9/244 (3.7%) | < 0.001*** |
| Low                                | 79/269 (29.4%) | 74/244 (30.3%) | < 0.001*** |
| Moderate                            | 46/269 (17.1%) | 115/244 (47.1%) | < 0.001*** |
| high                               | 10/269 (3.7%) | 46/244 (18.9%) | < 0.001*** |
| **Intralobular reticulations**     |            |            |         |
| No                                 | 231/269 (85.9%) | 103/244 (42.2%) | < 0.001*** |
| Yes                                | 38/269 (14.1%) | 141/244 (57.8%) |         |
| **Tree-in-bud micronodules**       |            |            |         |
| No                                 | 209/269 (77.7%) | 230/244 (94.3%) | < 0.001*** |
| Yes                                | 60/269 (22.3%) | 14/244 (5.7%) |         |
| **Presence of bronchial wall thickening** |        |            |         |
| No                                 | 194/269 (72.1%) | 227/244 (93%) | < 0.001*** |
| Yes                                | 75/269 (27.9%) | 17/244 (7%) |         |
| **Distribution of bronchial abnormalities** |        |            |         |
| None                               | 194/269 (72.1%) | 227/244 (93%) | < 0.001*** |
| Lobar/segmental                    | 29/269 (10.8%) | 7/244 (2.9%) | < 0.001*** |
| Diffuse                            | 46/269 (17.1%) | 10/244 (4.1%) | < 0.001*** |
| **Airway secretion**               |            |            |         |
| No                                 | 208/269 (77.3%) | 227/244 (93%) | < 0.001*** |
| Yes                                | 61/269 (22.7%) | 17/244 (7%) |         |
| **Fibrosis**                       |            |            |         |
| No                                 | 253/269 (94.1%) | 193/244 (79.1%) | < 0.001*** |
| Yes                                | 16/269 (5.9%) | 51/244 (20.9%) |         |
| **Pleural effusion**               |            |            |         |
|                          | Absent     | Yes, unilateral | Yes, bilateral | P-value |
|--------------------------|------------|-----------------|----------------|---------|
|                         | 221/269 (82.2%) | 18/269 (6.7%)   | 30/269 (11.2%) | 0.2     |
|                         | 215/244 (88.1%) | 10/244 (4.1%)   | 19/244 (7.8%)  | 0.3     |

| Adenomegaly              |            |                 |                |         |
|--------------------------|------------|-----------------|----------------|---------|
| No                       | 248/269 (92.2%) | 218/244 (89.3%) | 0.3            |         |
| Yes                      | 21/269 (7.8%)  | 26/244 (10.7%)  |                |         |

NOTE. – §. Data refer to number of patients with percentage in parentheses.

*: p < 0.05, **: p < 0.005, ***: p < 0.001 (corresponds to $X_2$ or Fisher test)

P-values in bold correspond to significant associations between clinical or radiological features and RT-PCR status. Regarding variables with more than two levels, the p value in italics corresponds to the p-value considering all its levels. The p-values below are based on the Fisher or $X_2$ test for this variable that was dichotomized according to the level of the line. Abbreviations: GGO: ground-glass opacities

**Table 3.** Final multivariate stepwise binary logistic regression models elaborated with either radiological or clinical-radiological variables.
| Variables§ | Coefficients (βi) | P-value | Multivariate OR (95% CI) | Examples §§ |
|------------|-------------------|---------|--------------------------|-------------|
|            |                   |         |                          | A    | B    | C    |
| **RADIOLOGICAL MODEL** |                   |         |                          |     |     |     |
| 0. (Intercept) | -2.758126351      | -       | 0.06 (0.03 - 0.12)       | 1    | 1    | 1    |
| 1. Presence of GGO | 1.089578460      | 0.04*   | 2.97 (1.03 - 8.56)       | 0    | 1    | 1    |
| 2. Fibrotic band | 1.5310410593     | <0.001*** | 4.62 (2.11 - 10.77)    | 1    | 0    | 1    |
| 3. GGO predominant pattern | 1.5366549953     | <0.001*** | 4.65 (1.96 - 11.45)    | 0    | 1    | 1    |
| 4. Subpleural predominant distribution | 0.9053515712   | 0.004** | 2.47 (1.33 - 4.63)       | 0    | 1    | 1    |
| 5. Diffuse lesions | 0.8414110324     | 0.02*   | 2.32 (1.16 - 4.61)       | 0    | 0    | 1    |
| 6. Intralobular reticulations | 0.7052466524     | 0.03*   | 2.02 (1.05 - 3.91)       | 0    | 0    | 1    |
| 7. Bronchial wall thickening | -1.758954141     | <0.001*** | 0.17 (0.07 - 0.4)        | 1    | 0    | 0    |
| **Probability for RT-PCR+§§** |                   |         |                          | 4.8% | 68.4% | 97.9% |
| **CLINICAL-RADIOLOGICAL MODEL** |                   |         |                          |     |     |     |
| 0. (Intercept) | -4.805582918      | -       | -                        | 1    | 1    | 1    |
| 1. Fever | 1.982642859       | <0.001*** | 7.26 (2.82 - 20.41)    | 1    | 1    | 1    |
| 2. Myalgia | 0.899937144       | 0.08    | 2.46 (0.82 - 6.94)       | 1    | 0    | 0    |
| 3. Presence of GGO | 1.181705996      | 0.08    | 3.26 (0.86 - 12.54)      | 1    | 1    | 0    |
| 4. Presence of fibrotic band | 1.322814497     | 0.02*   | 3.75 (1.34 - 11.47)      | 1    | 0    | 1    |
| 5. GGO predominant pattern | 1.504660153     | 0.008*  | 4.5 (1.52 - 14.24)       | 1    | 1    | 0    |
| 6. Peripheral predominant distribution | 0.930711612    | 0.03*   | 2.54 (1.12 - 5.83)       | 1    | 1    | 0    |
| 7. Diffuse lesions | 1.413720338      | 0.003** | 4.11 (1.66 - 10.58)      | 1    | 0    | 0    |
| 8. Intralobular reticulations | 1.028166435     | 0.02*   | 2.8 (1.2 - 6.66)         | 1    | 0    | 0    |
| 9. Bronchial wall thickening | -1.716180719     | 0.002** | 0.18 (0.06 - 0.51)       | 0    | 0    | 1    |
| **Probability for RT-PCR+§§§** |                   |         |                          | 99.6% | 68.9% | 3.9% |

**NOTE.** – §. The variables correspond to those that were kept in the final model after the stepwise backward-forward selection.

- §. Examples correspond to 6 distinct clinical cases. Each variable has 2 levels: “1” if the variable Xi is present (for instance fever), and “0” if the variable Xi is absent (for instance lack of fever). Hence, in example (A) with
the clinical-radiological model, the patient showed fever, myalgia, GGO, fibrotic bands, a GGO predominant pattern, a peripheral predominant distribution for abnormal acute findings and intralobular reticulations but no bronchial wall thickening.

- §§. The probability for RT-PCR+ are calculated as follows:

\[
P(\text{RT-PCR}+) = \frac{1}{1 + \exp[-(\beta_0 + \sum_{i=1}^{n} \beta_i \times X_i)]}
\]

*: \(P < .05\), **: \(P < .005\), ***: \(P < .001\). Significant results are highlighted in bold.

Abbreviations: 95%CI: 95% confidence interval; GGO: ground glass opacity; OR: odds ratio

Table 4. Performances of the final, trained, multivariate models and on-call teleradiologists’ conclusions on the independent validation cohort

| Performance measures | Final Models | OCTRs (validation cohort) |
|-----------------------|--------------|----------------------------|
|                       | Step-LR      | CART                       |                             |
| Radiological models (n = 101) |              |                            |                             |
| Accuracy§              | 0.83 (0.74 - 0.90) | 0.86 (0.78 - 0.92) | 0.81 (0.72 - 0.88) |
| OR§                   | 26.3 (8.9 - 77.9) | 56.1 (14.6 - 215.3) | 26.9 (8.2 - 88.4) |
| Sensitivity§           | 0.77 (0.63 - 0.88) | 0.77 (0.63 - 0.88) | 0.69 (0.54 - 0.81) |
| Specificity§           | 0.89 (0.77 - 0.96) | 0.94 (0.84 - 0.99) | 0.92 (0.82 - 0.98) |
| PPV§                  | 0.86 (0.74 - 0.93) | 0.93 (0.80 - 0.97) | 0.89 (0.76 - 0.96) |
| NPV§                  | 0.81 (0.72 - 0.88) | 0.82 (0.73 - 0.88) | 0.77 (0.68 - 0.83) |
| AUC                   | 0.90 (0.84 - 0.96) | 0.92 (0.86 - 0.98) | 0.88 (0.81 - 0.94) |

Clinical-radiological models (n = 71)§§

| Performance measures | Final Models | OCTRs (validation cohort) |
|-----------------------|--------------|----------------------------|
|                       | Step-LR      | CART                       |                             |
| Accuracy§              | 0.86 (0.76 - 0.93) | 0.83 (0.72 - 0.91) | 0.79 (0.68 - 0.88) |
| OR§                   | 44.2 (10.4 - 187.0) | 27.1 (7.4 - 100.) | 21.8 (5.5 - 85.5) |
| Sensitivity§           | 0.82 (0.66 - 0.92) | 0.80 (0.63 - 0.91) | 0.69 (0.52 - 0.83) |
| Specificity§           | 0.91 (0.75 - 0.98) | 0.88 (0.71 - 0.965) | 0.91 (0.75 - 0.98) |
| PPV§                  | 0.89 (0.73 - 0.96) | 0.85 (0.70 - 0.94) | 0.87 (0.69 - 0.95) |
| NPV§                  | 0.85 (0.74 - 0.92) | 0.82 (0.71 - 0.90) | 0.76 (0.67 - 0.84) |
| AUC                   | 0.93 (0.87 - 0.98) | 0.90 (0.83 - 0.97) | 0.86 (0.78 - 0.95) |
NOTE. – §. The diagnostic performance measures are calculated after dichotomizing the predicted probability for RT-PCR+ per 0.5 (i.e. < 0.5 corresponds to RT-PCR- and > 0.5 to RT-PCR+). Regarding the teleradiologists’ performances, they correspond to conclusion = (1), (2), (3) or (4) versus conclusion = (5).

- §. Thirty of the 101 patients from the validation cohort were excluded because of missing values (all from clinical variables).

Accuracy, OR, sensitivity, specificity, PPV, NPV and AUC are given with 95% confidence interval.

The highest diagnostic performance measure for each line is highlighted in bold.

Abbreviations: AUC: area under the ROC curve, CART: classification and regression tree; NPV: negative predictive value; OCTRs: on call teleradiologists; OR: odds ratio; PPV: positive predictive value; Step-LR: stepwise binary logistic regression.