



COVID-19 pneumonia accurately detected on chest radiographs with artificial intelligence



Francisco Dorr^a, Hernán Chaves^{a,b}, María Mercedes Serra^{a,b}, Andrés Ramirez^a,
Martín Elías Costa^a, Joaquín Seia^a, Claudia Cejas^b, Marcelo Castro^c, Eduardo Eyheremendy^d,
Diego Fernández Slezak^{a,e,f}, Mauricio F. Farez^{a,g,*}, on behalf of the Study Collaborators

^a Entelai, Buenos Aires, Argentina

^b Department of Diagnostic Imaging, Fleni, Buenos Aires, Argentina

^c Department of Diagnostic Imaging, Clínica Indisa, Santiago, Chile

^d Department of Diagnostic Imaging, Hospital Alemán, Buenos Aires, Argentina

^e Departamento de Computación, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Buenos Aires, Argentina

^f Instituto de Investigación en Ciencias de la Computación (ICC), CONICET-Universidad de Buenos Aires, Buenos Aires, Argentina

^g Center for Epidemiology, Biostatistics and Public Health, Fleni, Buenos Aires, Argentina

ARTICLE INFO

Keywords:
Chest
Radiography
Artificial intelligence
Diagnostic performance
COVID-19

ABSTRACT

Purpose: To investigate the diagnostic performance of an Artificial Intelligence (AI) system for detection of COVID-19 in chest radiographs (CXR), and compare results to those of physicians working alone, or with AI support.

Materials and methods: An AI system was fine-tuned to discriminate confirmed COVID-19 pneumonia, from other viral and bacterial pneumonia and non-pneumonia patients and used to review 302 CXR images from adult patients retrospectively sourced from nine different databases. Fifty-four physicians blind to diagnosis, were invited to interpret images under identical conditions in a test set, and randomly assigned either to receive or not receive support from the AI system. Comparisons were then made between diagnostic performance of physicians working with and without AI support. AI system performance was evaluated using the area under the receiver operating characteristic (AUROC), and sensitivity and specificity of physician performance compared to that of the AI system.

Results: Discrimination by the AI system of COVID-19 pneumonia showed an AUROC curve of 0.96 in the validation and 0.83 in the external test set, respectively. The AI system outperformed physicians in the AUROC overall (70% increase in sensitivity and 1% increase in specificity, $p < 0.0001$). When working with AI support, physicians increased their diagnostic sensitivity from 47% to 61% ($p < 0.001$), although specificity decreased from 79% to 75% ($p = 0.007$).

Conclusions: Our results suggest interpreting chest radiographs (CXR) supported by AI, increases physician diagnostic sensitivity for COVID-19 detection. This approach involving a human-machine partnership may help expedite triaging efforts and improve resource allocation in the current crisis.

1. Introduction

Starting on December 8, 2019, a series of viral pneumonia cases of unknown etiology emerged in Wuhan, Hubei province, China [1–3]. Sequencing analysis from respiratory tract samples identified a novel coronavirus, tentatively named 2019-nCoV by the World Health Organization and subsequently designated as SARS-CoV-2 by the

International Committee on Taxonomy of Viruses [4]. During the first two months of 2020, the virus causing the disease known as COVID-19 spread worldwide, showing evidence of human-to-human transmission between close contacts [5]. The World Health Organization declared the coronavirus outbreak a pandemic on March 11, and countries around the world struggled with an unprecedented surge in confirmed cases [6]. SARS-CoV-2 causes varying degrees of illness, the most common

Abbreviations: RT-PCR, real-time reverse transcriptase–polymerase chain reaction; CXR, chest radiographs; DL, deep learning; AUROC, area under the receiver operating characteristic; AUPR, area under the precision-recall; CT, computed tomography; AI, artificial intelligence.

* Corresponding author. Montañeses 2325, 7th Floor, Buenos Aires, 1428, Argentina, ;

E-mail addresses: mfarez@entelai.com, mfarez@fleni.org.ar (M.F. Farez).

<https://doi.org/10.1016/j.ibmed.2020.100014>

Received 10 July 2020; Received in revised form 6 November 2020; Accepted 12 November 2020

2666-5212/© 2020 Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

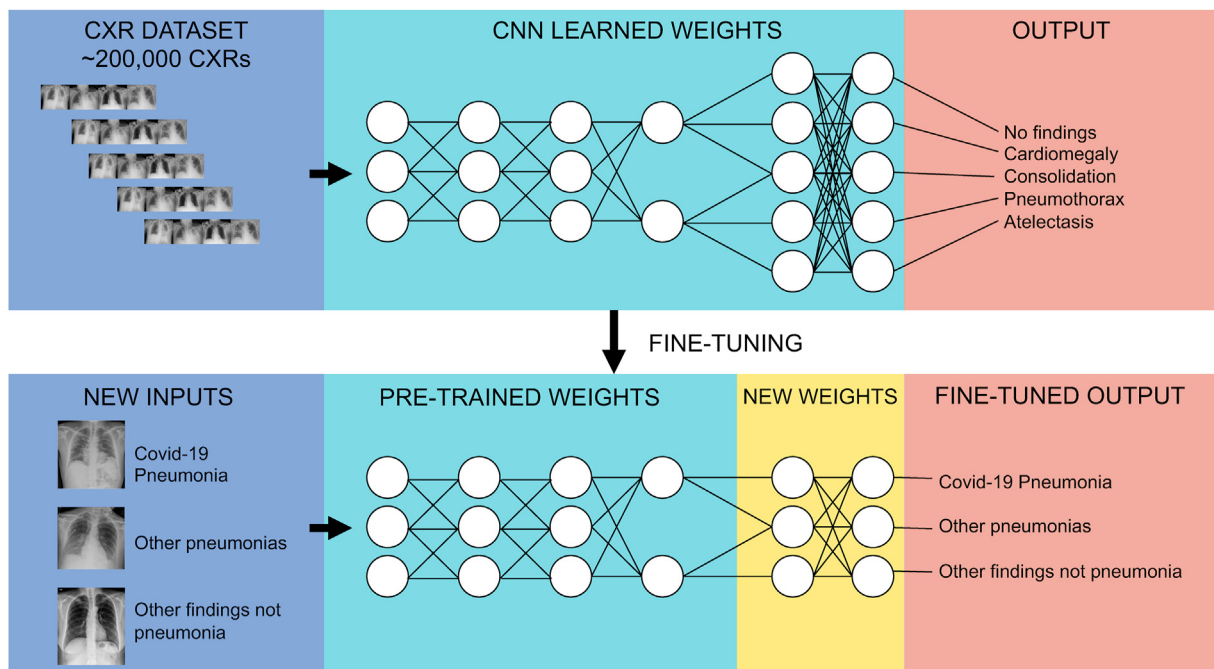


Fig. 1. Convolutional Neural Network Diagram. This chart summarizes the strategy used in the study. Using a convolutional neural network, pre-trained with a dataset of over 200,000 CXRs and 5 output classes; all layers but the last block of layers were frozen and transferred onto a new network with new labels (COVID-19 pneumonia, Other pneumonias, Normal/Other findings). Final fully-connected layers were then retrained over the transferred ones.

release. However, non-privative code parts have been released in a public repository that can be found in <https://bitbucket.org/aenti/entelai-vid-paper>. All study experiments and implementation methods are described in detail and the tool itself is available online at: <https://covid.entelai.com>, to enable independent replication.

2.6. Data availability

Local datasets and links to image repositories used in the study are publicly available online.²

3. Results

3.1. Training and validation of the AI system

We fine-tuned a pre-established AI system using a dataset of 302 CXR of COVID-19, other pneumonia, and other non-pneumonia cases. After 20 epochs of training, we obtained a mean AUROC curve among the 5 cross-validation folds of 0.96 ± 0.02 (see Fig. 2 and Table 1).

One of the traditional criticisms of DL models is the risk of "black box" predictions, implying the information that the model uses to make predictions is unclear and may not be meaningful. Recently, activation maps have been developed as a way to depict what the models are using to support their predictions [29]. We analyzed activation maps for COVID-19 and compared them to other pneumonias, to validate the model and identify potential sources of information. The activation maps were obtained by taking the output of the average pooling layer and taking the mean across channel dimension [30]. As shown in Fig. 3, activation maps generated using this AI system relied heavily on lower pulmonary lobes, and on peripheral lung regions in particular. Of note, peripheral infection patterns have recently been described as a key feature in COVID-19 [8,31], suggesting the AI system was able to predict COVID-19 diagnosis using relevant information from CXRs.

² https://osf.io/6by7h/?view_only=28264f73003245f897a847d3cd496ab9.

Since training can overfit prediction to a particular dataset, we generated an independent test set comprising 60 images (20 per category) to evaluate AI system performance. AUROC, Brier and Mean Absolute Error scores were obtained on a one-vs-rest basis. Brier scores in particular are widely used in medical research to assess and compare model prediction accuracy [32]. Values range from 0 to 1, with 0 being the best possible outcome. Although they can be used as a single multi-class score, in this study we reported Brier scores by class, to obtain a better idea of how well the model performed for each one. As shown in Table 2 and Fig. 4, performance of the model was not as good, but nevertheless acceptable, since this AI system was able to predict COVID-19 with a sensitivity and specificity of 80% and an AUROC of 0.84. This difference between the cross-validation and the test results could be explained by the data sets used. Since the number of instances of each dataset is low, it is almost impossible to obtain a perfect generalization. The model could be learning certain particularities of the training set that, in spite of doing a cross-validation and having regularization by dropout, the overfitting to the specific dataset could not be completely overcome. More data will be needed to achieve a similar score between the cross-validation and the test set.

3.2. Clinical performance results

We next analyzed whether identification and separation of COVID-19 by physicians was adequate, given the novelty of the disease and the lack of worldwide experience. To this end, we tested the performance of 60 physicians from several different referral centers in South America. Six physicians were excluded for not completing the survey in time [$n = 4$], or not answering a minimum number of questions [$n = 2$]. Fifty-four physicians from Argentina [$n = 49$], Chile [$n = 4$] and Colombia [$n = 1$] were included. Given the good performance of the model, we randomly informed physicians what the AI system prediction had been for 50% of the images (which could be correct or incorrect as per its performance on the same Test Set). AI system prediction was shared with physicians as a likelihood percentage for each condition. Physicians would then have to give the most likely diagnosis, given the AI suggestion. As shown in Fig. 4, sensitivity and specificity for COVID-19

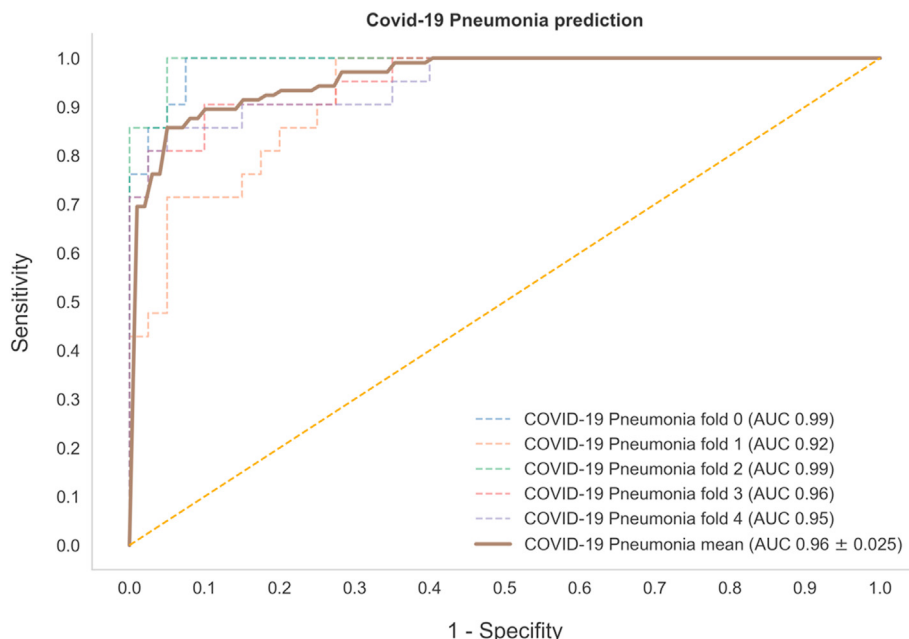


Fig. 2. Performance of the Artificial Intelligence (AI) System in COVID-19 Prediction. Receiver operating characteristic (ROC) curve and area under the curve (AUC) of the AI system on the validation set for each of the 5 folds, with a mean area under the receiver operating characteristic (AUROC) curve of 0.96 ± 0.02 , $n = 302$.

Table 1

Performance of the AI system in the training dataset using the average of 5-fold cross-validation.

Diagnosis	Sensitivity	Specificity	AUROC
Covid-19 pneumonia (n = 102)	94%	81%	0.96
Non-Covid-19 pneumonia (n = 100)	55%	95%	0.87
Other (n = 100)	84%	91%	0.93

prediction based on CXR by physicians was 47% and 79% respectively, with an increase in sensitivity to 61% ($p < 0.001$) and a decrease in specificity to 74% ($p = 0.007$) when using AI support. No significant differences between radiologists and emergency care physicians were observed, nor did years of training affect overall performance results (data not shown).

4. Discussion

In the setting of the COVID-19 pandemic, it is probable that RT-PCR tests will become more robust, quicker, and ubiquitous. However, due to

the actual shortage and limitations of RT-PCR kits, diagnostic imaging modalities such as CXR and CT have been proposed as surrogate methods for COVID-19 triage. Some researchers have even reported chest CTs as showing higher sensitivity for COVID-19 detection than RT-PCR from swab samples [33,34]. Mei et al. went further and used AI to integrate chest CT findings with clinical symptoms, exposure history and laboratory testing achieving an AUROC of 0.92 and had equal sensitivity as compared to a senior thoracic radiologist [35]. However, the American College of Radiology currently recommends CT be reserved for hospitalized, symptomatic patients with specific clinical indications [17]. CT also increases exposure to radiation, is less cost-effective, not widely available and requires appropriate infection control procedures during and after examination, including closing scanning rooms for up to 1 h for airborne precaution measures [36]. This is why CXR (the most commonly performed diagnostic imaging examination) has been proposed as first-line imaging study when COVID-19 is suspected, especially in resource-constrained scenarios [11,12]. Portable X-ray units are particularly suitable, as they can be moved to the emergency department (ED) or intensive care unit and easily cleaned afterwards [17].

Most clinicians have less experience interpreting CXRs than

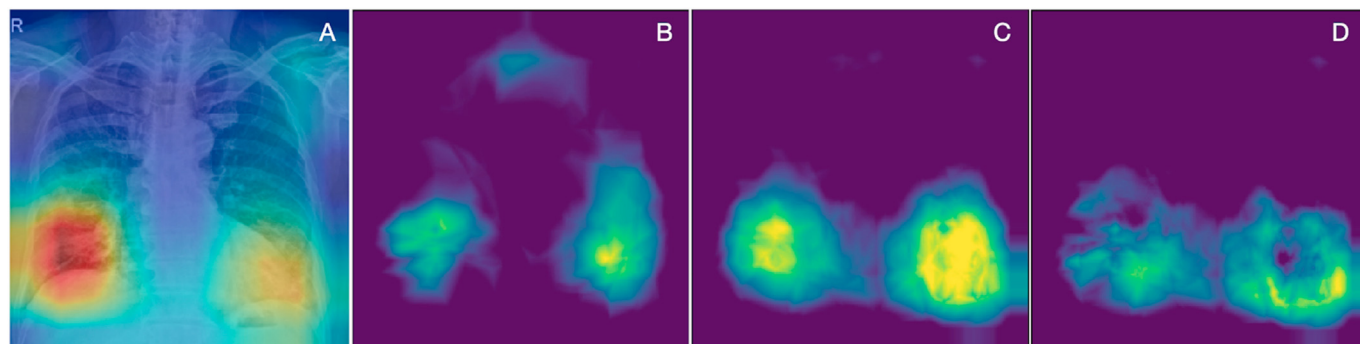


Fig. 3. Activation Maps of the Artificial Intelligence (AI) System. a) Example of a single activation map on a CXR image from the COVID-19 group. b) Mean activation map of Non-COVID-19 pneumonia category. c) Mean activation map of COVID-19 pneumonia category. d) Delta activation map between COVID-19 and Non-COVID-19 pneumonia categories calculated by $\max_{i,j}(CovidMeanMap_{i,j} - NonCovidMeanMap_{i,j}, 0)$ for each pixel (i,j) , depicting lower and peripheral areas as more relevant for the differentiation.

Table 2
Performance of the AI system in the test dataset.

Diagnosis	Sensitivity	Specificity	AUROC	F1 score	Brier score	MAE
Covid-19 pneumonia (n = 20)	80%	80%	0.84	0.73	0.16	0.28
Non-Covid-19 pneumonia (n = 20)	60%	90%	0.88	0.67	0.14	0.26
Other (n = 20)	65%	83%	0.86	0.65	0.15	0.26

AI: artificial intelligence, AUROC: area under the receiver operating characteristics, MAE: mean absolute error.

radiologists. In the ED setting however, physicians with no formal radiology training are the ones most often reporting CXR findings. Gatt et al., found sensitivity levels as low as 20% for CXR evaluation results by emergency care physicians [37]. One would expect this sensitivity to be even lower in the setting of a new disease like COVID-19. At the other end of the spectrum, Wong et al. found thoracic radiologist sensitivity level for CXR diagnosis in a cohort of COVID-19 patients was 69% at baseline [38], and Cozzi et al. found sensitivities as high as 100% in experienced radiologists [39]. In our study we noted a low sensibility (both in radiologist and emergency care physicians) for the diagnosis of COVID-19 pneumonia. This could be explained by the fact that, at the time of the clinical study, most physicians that participated in the survey, have been exposed to few COVID-19 cases. Low sensibility could also be related to the online survey design, as physicians evaluated CXR in a different fashion as they do in their clinical practice, with a limited amount of time to give a diagnosis. We also noted decreased specificity, due to increased numbers of false positives in the AI-supported group. In every case, false positives arose from doubts over the “Other Pneumonias” category; although the AI model correctly predicted and presented the label “Other Pneumonias”, physicians were still inclined to favor a COVID-19 diagnosis. The significance and clinical impact of this effect is unclear and deserves further evaluation.

AI has proven useful in CXR analysis for many diseases [18–22]. In the setting of COVID-19 emergence, several AI models based on DL have

been developed around the world, with varying results in terms of accuracy detecting COVID-19-infected patients based on CXR [23–26]. Moreover, none of these models have been tested in real or simulated clinical scenarios.

Murphy et al. developed an AI system for the evaluation of CXR in the setting of COVID-19 and achieved a lower AUROC (0.81), and their test set came from a single institution [40]. They compared the performance of the AI system to radiologist performance but did not evaluate the change in diagnostic accuracy of radiologist without and with AI support as we did.

Considering the prevalence of adults in the COVID-19 group, we chose to exclude pediatric databases to avoid major bias in training and testing.

Early diagnosis, isolation and prompt clinical management are the three public health strategies collectively contributing to contain the spread of COVID-19. AI models building on the first of these premises might be significant [41]. In this study, we designed and evaluated a DL model trained to detect COVID-19 on CXR images. On an independent test dataset, the model showed 80% sensitivity and specificity for COVID-19 detection with an AUROC value of 0.84. We also observed improved diagnostic sensitivity in physician performance (both for radiologists and emergency care physicians) and decreased specificity. Of note, despite AI system support, physicians did not reach or surpass AI metrics. Our results differ from the work of Patel et al. who tested a model in a simulated clinical scenario applied to CXR pneumonia diagnosis and achieved maximum diagnostic accuracy combining radiologist and AI performance [42]. This could have been due to lack of formal training to incorporate AI recommendations, or lack of trust in our model predictions. Both hypotheses should be further validated in future studies.

Our model has significant limitations. First, despite the large number of CXR used to train the original model (around 224,000 images), only a small number of CXRs were added to our DL model (around 100 images per category) using a transfer learning approach. Second, our training set is mostly based on adult patients CXRs from China and Italy. Third, our model could also be prone to selection bias, as databases tend to include more severe or complicated cases. Since the disease has emerged recently, few good quality, curated, COVID-19 CXR databases are

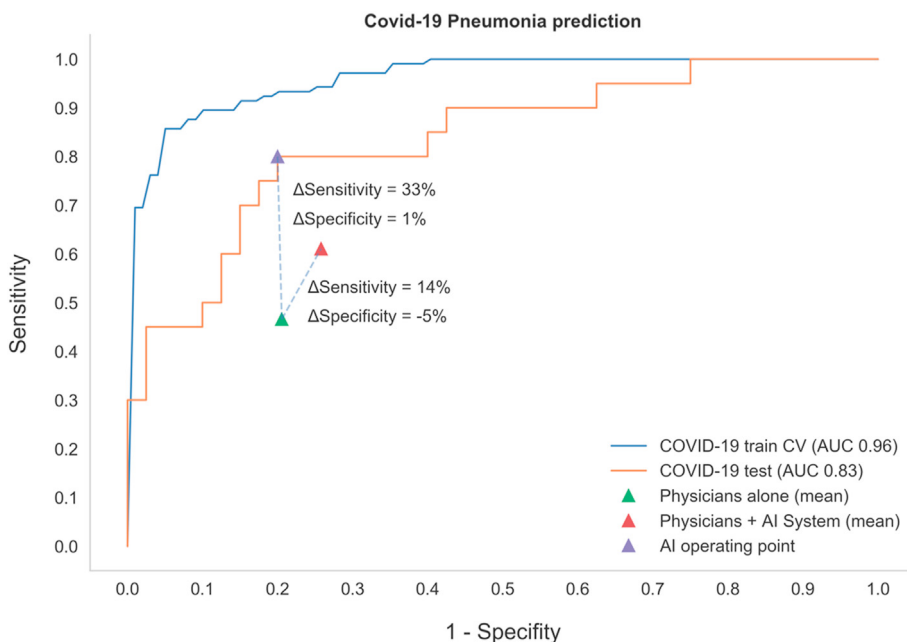


Fig. 4. Performance of the Artificial Intelligence (AI) System on the Train and Test Sets, Compared to the Performance of Physicians in COVID-19 Prediction. Receiver operating characteristic (ROC) curve and area under the curve (AUC) of the AI system on the train and test sets. Physician performance with and without AI support is compared.

available. Inclusion of cases of all ages, from every region around the world, would certainly improve AI systems diagnostic accuracy and reliability.

5. Conclusions

In conclusion, our data suggests physician performance can be improved using AI systems such as the one described here. We showed an increase in sensitivity from 47% to 61% for COVID-19 prediction based on CXR. Future prospective studies are needed to further evaluate the clinical and public health impact of the combined work of physicians and AI systems.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Mauricio F. Farez has received professional travel/accommodations stipends from Merck-Serono Argentina, Teva Argentina and Novartis Argentina. The rest of the authors declare no competing interests.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.ibmed.2020.100014>.

References

- Hui DS, I Azhar E, Madani TA, Ntoumi F, Kock R, Dar O, et al. The continuing 2019-nCoV epidemic threat of novel coronaviruses to global health — the latest 2019 novel coronavirus outbreak in Wuhan, China. *Int J Infect Dis* 2020;91:264–6. <https://doi.org/10.1016/j.ijid.2020.01.009>.
- Lu H, Stratton CW, Tang Y. Outbreak of pneumonia of unknown etiology in Wuhan, China: the mystery and the miracle. *J Med Virol* 2020;92:401–2. <https://doi.org/10.1002/jmv.25678>.
- Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* 2020;395:497–506. [https://doi.org/10.1016/S0140-6736\(20\)30183-5](https://doi.org/10.1016/S0140-6736(20)30183-5).
- Gorbalenya AE, Baker SC, Baric RS, de Groot RJ, Drosten C, Gulyaeva AA, et al. Severe acute respiratory syndrome-related coronavirus: the species and its viruses – a statement of the Coronavirus Study Group. *Microbiology* 2020. <https://doi.org/10.1101/2020.02.07.937862>.
- Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, et al. Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. *N Engl J Med* 2020;382:1199–207. <https://doi.org/10.1056/NEJMoa2001316>.
- World Health Organization. Coronavirus disease (COVID-19) outbreak. <https://www.who.int/emergencies/diseases/novel-coronavirus-2019>. [Accessed 12 April 2020].
- Chen N, Zhou M, Dong X, Qu J, Gong F, Han Y, et al. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *Lancet* 2020;395:507–13. [https://doi.org/10.1016/S0140-6736\(20\)30211-7](https://doi.org/10.1016/S0140-6736(20)30211-7).
- Guan W, Ni Z, Hu Y, Liang W, Ou C, He J, et al. Clinical characteristics of coronavirus disease 2019 in China. *N Engl J Med* 2020. <https://doi.org/10.1056/NEJMoa2002032>.
- Ministerio de Salud de la Republica Argentina. Definición de caso. <https://www.argentina.gob.ar/salud/coronavirus-COVID-19/definicion-de-caso>. [Accessed 19 April 2020].
- Ministerio de Sanidad de España. Manejo en urgencias del COVID-19. https://www.mscbs.gob.es/profesionales/saludPublica/ccayes/alertasActual/nCov-China/documentos/Manejo_urgencias_pacientes_con_COVID-19.pdf. [Accessed 20 April 2020].
- The British society of thoracic imaging, BSTI NHSE COVID-19 radiology decision support tool. <https://www.bsti.org.uk/standards-clinical-guidelines/clinical-guidelines/bsti-nhse-covid-19-radiology-decision-support-tool/>. [Accessed 21 April 2020].
- Rubin GD, Ryerson CJ, Haramati LB, Sverzellati N, Kanne JP, Raoof S, et al. The role of chest imaging in patient management during the COVID-19 pandemic: a multinational consensus statement from the Fleischner society. *Chest* 2020;158:106–16. <https://doi.org/10.1016/j.chest.2020.04.003>.
- World Health Organization. Coronavirus disease (COVID-19) technical guidance: laboratory testing for 2019-nCoV in humans. <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/laboratory-guidance>. [Accessed 12 April 2020].
- American Society for Microbiology. ASM expresses concern about coronavirus test reagent shortages. <https://asm.org/Articles/Policy/2020/March/ASM-Expresses-Concern-about-Test-Reagent-Shortages>. [Accessed 12 April 2020].
- The New Yorker. Why widespread coronavirus testing isn't coming anytime soon. <https://www.newyorker.com/news/news-desk/why-widespread-coronavirus-testing-isnt-coming-anytime-soon>. [Accessed 12 April 2020].
- Yang Y, Yang M, Shen C, Wang F, Yuan J, Li J, et al. Evaluating the accuracy of different respiratory specimens in the laboratory diagnosis and monitoring the viral shedding of 2019-nCoV infections. *Infectious Diseases (except HIV/AIDS)* 2020. <https://doi.org/10.1101/2020.02.11.20021493>.
- American College of Radiology. ACR recommendations for the use of chest radiography and computed tomography (CT) for suspected COVID-19 infection. <https://www.acr.org/Advocacy-and-Economics/ACR-Position-Statements/Recommendations-for-Chest-Radiography-and-CT-for-Suspected-COVID19-Infection>. [Accessed 16 April 2020].
- Singh R, Kalra MK, Nitiwarangkul C, Patti JA, Homayounieh F, Padole A, et al. Deep learning in chest radiography: detection of findings and presence of change. *PLoS One* 2018;13. <https://doi.org/10.1371/journal.pone.0204155>.
- Rajpurkar P, Irvin J, Ball RL, Zhu K, Yang B, Mehta H, et al. Deep learning for chest radiograph diagnosis: a retrospective comparison of the CheXNeXt algorithm to practicing radiologists. *PLoS Med* 2018;15:e1002686. <https://doi.org/10.1371/journal.pmed.1002686>.
- Lee SM, Seo JB, Yun J, Cho Y-H, Vogel-Claussen J, Schiebler ML, et al. Deep learning applications in chest radiography and computed tomography: current state of the art. *J Thorac Imag* 2019;34:75–85. <https://doi.org/10.1097/RTI.0000000000000387>.
- Kermany DS, Goldbaum M, Cai W, Valentim CCS, Liang H, Baxter SL, et al. Identifying medical diagnoses and treatable diseases by image-based deep learning. *Cell* 2018;172:1122–31. <https://doi.org/10.1016/j.cell.2018.02.010>. e9.
- Kallianos K, Mongan J, Antani S, Henry T, Taylor A, Abuaya J, et al. How far have we come? Artificial intelligence for chest radiograph interpretation. *Clin Radiol* 2019;74:338–45. <https://doi.org/10.1016/j.crad.2018.12.015>.
- Wang L, Wong A. COVID-net: a tailored deep convolutional neural network design for detection of COVID-19 cases from chest radiography images. 2020. ArXiv:200309871 [Cs, Eess].
- Khan AI, Shah JL, CoroNet Bhat M. A deep neural network for detection and diagnosis of covid-19 from chest X-ray images. 2020. ArXiv:200404931 [Cs, Eess, Stat].
- Hemdan EE-D, Shouman MA, Karar ME. COVIDX-net: a framework of deep learning classifiers to diagnose COVID-19 in X-ray images. 2020. ArXiv:200311055 [Cs, Eess].
- Afshar P, Heidarian S, Naderkhani F, Oikonomou A, Plataniotis KN, Mohammadi A. COVID-CAPS: a capsule network-based framework for identification of COVID-19 cases from X-ray images. 2020. ArXiv:200402696 [Cs, Eess].
- Irvin J, Rajpurkar P, Ko M, Yu Y, Ciurea-Illcus S, Chute C, et al. CheXpert: a large chest radiograph dataset with uncertainty labels and expert comparison. 2019. ArXiv:190107031 [Cs, Eess].
- Huang G, Liu Z, van der Maaten L, Weinberger KQ. Densely connected convolutional networks. 2018. ArXiv:160806993 [Cs].
- Arrieta AB, Díaz-Rodríguez N, Del Ser J, Bennetot A, Tabik S, Barbado A, et al. Explainable artificial intelligence (XAI): concepts, taxonomies, opportunities and challenges toward responsible AI. 2019. ArXiv:191010045 [Cs].
- Zhou B, Khosla A, Lapedriza A, Oliva A, Torralba A. Learning deep features for discriminative localization. 2015. ArXiv:151204150 [Cs].
- Salehi S, Abedi A, Balakrishnan S, Gholamrezaehezad A. Coronavirus disease 2019 (COVID-19): a systematic review of imaging findings in 919 patients. *Am J Roentgenol* 2020;1–7. <https://doi.org/10.2214/AJR.20.23034>.
- Ruffbach K. Use of Brier score to assess binary predictions. *J Clin Epidemiol* 2010;63:938–9. <https://doi.org/10.1016/j.jclinepi.2009.11.009>.
- Ai T, Yang Z, Hou H, Zhan C, Chen C, Lv W, et al. Correlation of chest CT and RT-PCR testing in coronavirus disease 2019 (COVID-19) in China: a report of 1014 cases. *Radiology* 2020;200642. <https://doi.org/10.1148/radiol.2020200642>.
- Fang Y, Zhang H, Xie J, Lin M, Ying L, Pang P, et al. Sensitivity of chest CT for COVID-19: comparison to RT-PCR. *Radiology* 2020;200432. <https://doi.org/10.1148/radiol.2020200432>.
- Mei X, Lee H-C, Diao K-Y, Huang M, Lin B, Liu C, et al. Artificial intelligence-enabled rapid diagnosis of patients with COVID-19. *Nat Med* 2020;26:1224–8. <https://doi.org/10.1038/s41591-020-0931-3>.
- Mossa-Basha M, Medverd J, Linnau K, Lynch JB, Wener MH, Kicska G, et al. Policies and guidelines for COVID-19 preparedness: experiences from the University of Washington. *Radiology*; 2020, 201326. <https://doi.org/10.1148/radiol.2020201326>.
- Gatt ME. Chest radiographs in the emergency department: is the radiologist really necessary? *Postgrad Med* 2003;79:214–7. <https://doi.org/10.1136/pmj.79.930.214>.
- Wong HYF, Lam HYS, Fong AH-T, Leung ST, Chin TW-Y, Lo CSY, et al. Frequency and distribution of chest radiographic findings in COVID-19 positive patients. *Radiology* 2020;201160. <https://doi.org/10.1148/radiol.2020201160>.
- Cozzi A, Schiaffino S, Arpaia F, Della Pepa G, Tritella S, Bertolotti P, et al. Chest x-ray in the COVID-19 pandemic: radiologists' real-world reader performance. *Eur J Radiol* 2020;132:109272. <https://doi.org/10.1016/j.ejrad.2020.109272>.
- Murphy K, Smits H, Knoops AJG, Korst MBJM, Samson T, Scholten ET, et al. COVID-19 on chest radiographs: a multireader evaluation of an artificial intelligence

- system. *Radiology* 2020;296:E166–72. <https://doi.org/10.1148/radiol.2020201874>.
- [41] Vaishya R, Javaid M, Khan IH, Haleem A. Artificial Intelligence (AI) applications for COVID-19 pandemic. *Diabetes & Metabolic Syndrome: Clin Res Rev* 2020;14:337–9. <https://doi.org/10.1016/j.dsx.2020.04.012>.
- [42] Patel BN, Rosenberg L, Willcox G, Baltaxe D, Lyons M, Irvin J, et al. Human-machine partnership with artificial intelligence for chest radiograph diagnosis. *Npj Digital Medicine* 2019;2:1–10. <https://doi.org/10.1038/s41746-019-0189-7>.