Radiology

Diagnosis of Coronavirus Disease 2019 Pneumonia by Using Chest Radiography: Value of Artificial Intelligence

Ran Zhang, PhD • Xin Tie, BS • Zhihua Qi, PhD • Nicholas B. Bevins, PhD • Chengzhu Zhang, MS • Dalton Griner, MS • Thomas K. Song, MD • Jeffrey D. Nadig, MD • Mark L. Schiebler, MD • John W. Garrett, PhD • Ke Li, PhD • Scott B. Reeder, MD, PhD • Guang-Hong Chen, PhD

From the Departments of Medical Physics (R.Z., X.T., C.Z., D.G., J.W.G., K.L., S.B.R., G.H.C.) and Radiology (M.L.S., J.W.G., K.L., S.B.R., G.H.C.), University of Wisconsin–Madison School of Medicine and Public Health, 1111 Highland Ave, Madison, WI 53705; and Department of Radiology, Henry Ford Health System, Detroit, Mich (Z.Q., N.B.B., T.K.S., J.D.N,). Received July 1, 2020; revision requested July 28; revision received September 15; accepted September 17. **Address correspondence to** G.H.C. (e-mail: *gchen7@wisc.edu*).

Study supported by a grant from Wisconsin Partnership Program and by the National Institute of Biomedical Imaging and Bioengineering (grant 3U01EB021183W4).

Conflicts of interest are listed at the end of this article.

Radiology 2021; 298:E88–E97 • https://doi.org/10.1148/radiol.2020202944 • Content code: IN

Background: Radiologists are proficient in differentiating between chest radiographs with and without symptoms of pneumonia but have found it more challenging to differentiate coronavirus disease 2019 (COVID-19) pneumonia from non–COVID-19 pneumonia on chest radiographs.

Purpose: To develop an artificial intelligence algorithm to differentiate COVID-19 pneumonia from other causes of abnormalities at chest radiography.

Materials and Methods: In this retrospective study, a deep neural network, CV19-Net, was trained, validated, and tested on chest radiographs in patients with and without COVID-19 pneumonia. For the chest radiographs positive for COVID-19, patients with reverse transcription polymerase chain reaction results positive for severe acute respiratory syndrome coronavirus 2 with findings positive for pneumonia between February 1, 2020, and May 30, 2020, were included. For the non–COVID-19 chest radiographs, patients with pneumonia who underwent chest radiography between October 1, 2019, and December 31, 2019, were included. Area under the receiver operating characteristic curve (AUC), sensitivity, and specificity were calculated to characterize diagnostic performance. To benchmark the performance of CV19-Net, a randomly sampled test data set composed of 500 chest radiographs in 500 patients was evaluated by the CV19-Net and three experienced thoracic radiologists.

Results: A total of 2060 patients (5806 chest radiographs; mean age, 62 years ± 16 [standard deviation]; 1059 men) with COVID-19 pneumonia and 3148 patients (5300 chest radiographs; mean age, 64 years ± 18; 1578 men) with non–COVID-19 pneumonia were included and split into training and validation and test data sets. For the test set, CV19-Net achieved an AUC of 0.92 (95% CI: 0.91, 0.93). This corresponded to a sensitivity of 88% (95% CI: 87, 89) and a specificity of 79% (95% CI: 77, 80) by using a high-sensitivity operating threshold, or a sensitivity of 78% (95% CI: 77, 79) and a specificity of 89% (95% CI: 88, 90) by using a high-specificity operating threshold. For the 500 sampled chest radiographs, CV19-Net achieved an AUC of 0.94 (95% CI: 0.93, 0.96) compared with an AUC of 0.85 (95% CI: 0.81, 0.88) achieved by radiologists.

Conclusion: CV19-Net was able to differentiate coronavirus disease 2019–related pneumonia from other types of pneumonia, with performance exceeding that of experienced thoracic radiologists.

Online supplemental material is available for this article.

The outbreak of coronavirus disease 2019 (COVID-19) (1) began with the initial diagnosis of an unknown viral pneumonia in late 2019 in Wuhan, China, that subsequently spread around the globe as a pandemic. Ribonucleic acid sequencing of respiratory samples identified a novel coronavirus (called severe acute respiratory syndrome coronavirus 2) as the underlying cause of COVID-19. Patients with COVID-19 present with symptoms that are similar to other viral illnesses including influenza and other coronaviruses such as severe acute respiratory syndrome (2,3) and Middle East respiratory syndrome (4). Symptoms are nonspecific and include fever, cough, fatigue, dyspnea, diarrhea, and even anosmia (5,6). The radiographic signs are also nonspecific and can be observed in patients with other viral illnesses, drug reactions, or aspiration (5,7,8).

The similarities in clinical presentation across other reactions and illnesses creates challenges regarding establishment of a clinical diagnosis. Currently, reverse transcription polymerase chain reaction (RT-PCR) is the reference standard method to identify patients with COVID-19 infection (9). In addition to the RT-PCR test, CT has also been widely used in China, and occasionally in other countries, to provide additional means of COVID-19 diagnosis and treatment-response monitoring (5,10,11). However, because of concerns of CT facility contamination and exposure to health care workers, health care professional organizations (12–14) do not recommend CT imaging as a general diagnostic imaging tool for patients with COVID-19.

Major medical societies instead recommend the use of chest radiography as part of the work-up for persons

Abbreviations

AUC = area under the receiver operating characteristic curve, COVID-19 = coronavirus disease 2019, RT-PCR = reverse transcription polymerase chain reaction

Summary

An artificial intelligence algorithm differentiated between coronavirus disease 2019 (COVID-19) pneumonia and non–COVID-19 pneumonia on chest radiographs with high sensitivity and specificity.

Key Results

- The overall performance of artificial intelligence algorithm achieved an area under the receiver operating characteristic curve (AUC) of 0.92 on the test data set of 5869 chest radiographs in 2193 patients (acquired from multiple hospitals and multiple vendors).
- On a set of 500 randomly selected test chest radiographs, the artificial intelligence algorithm achieved an AUC of 0.94, compared with an AUC of 0.85 from three experienced thoracic radiologists.

suspected of having COVID-19 because of its unique advantages: Almost all clinics, emergency departments, urgent care facilities, and hospitals are equipped with stationary and mobile radiography units, including both urban and rural medical facilities. These units can be easily protected from exposure or disinfected after use and can be directly used in a contained clinical environment without having to move patients. However, the major challenge with the use of chest radiography in diagnosis of COVID-19 is its low sensitivity and specificity in current radiologic practice. A recent study found that the sensitivity of chest radiography was poor for COVID-19 diagnosis (11). To some extent, this poor diagnostic performance can be attributed to the fact that many radiologists are observing COVID-19–induced pneumonia for the first time and therefore need to interpret more images to learn both the common and unique imaging features of this disease.

Machine learning methods, particularly deep learning (15,16), have unique advantages in quick and tireless learning to differentiate COVID-19 pneumonia from other types of pneumonia by using chest radiographs. The purpose of our study was to train and validate a deep learning method to differentiate COVID-19 pneumonia from other causes of abnormalities at chest radiography and to test its performance against thoracic radiologists.

Materials and Methods

This retrospective Health Insurance Portability and Accountability Act–compliant study was approved by the institutional review boards at Henry Ford Health System (Detroit, Mich) and the University of Wisconsin–Madison (Madison, Wis). Written informed consent was waived because of the retrospective nature of the data collection and the use of deidentified images.

Patient Data Sets

For algorithm development, we included chest radiographs in patients with and without COVID-19 pneumonia from Henry Ford Health System, which includes five hospitals and more than 30 clinics. The pneumonia findings for both COVID-19 and non–COVID-19 pneumonia were found by using a commercial natural language processing tool (InSight; Softek Illuminate) that searched radiologist reports for positive findings. Searches were performed in all radiologist reports at the institution during the COVID-19 and non–COVID-19 periods. The patients with non–COVID-19 pneumonia were selected solely on the basis of findings positive for pneumonia in the report and the date of study (October to December 2019). The patients with pneumonia from the COVID-19 period were cross-referenced with the list of patients positive for COVID-19 to find a list of patients with results positive for pneumonia and positive for COVID-19.

The inclusion criteria for the group with non–COVID-19 pneumonia were patients who underwent frontal view chest radiography, were diagnosed with pneumonia, and underwent imaging between October 1, 2019, and December 31, 2019 (the first patient who tested positive for COVID-19 in the United States was confirmed on January 19, 2020, in Seattle, Wash [17]). Because these chest radiographs predate the first confirmed cases of COVID-19 in the United States, we consider these chest radiographs to be positive for non–COVID-19 pneumonia. Patients younger than 18 years were excluded.

The inclusion criteria for the COVID-19–positive group were patients who underwent frontal view chest radiography, and who had an RT-PCR test result positive for severe acute respiratory syndrome coronavirus 2 with a diagnosis of pneumonia between February 1, 2020, and May 31, 2020. Patients were excluded if chest radiography was performed more than 5 days before or 14 days after RT-PCR confirmation.

The resulting data sets consisted of 5805 chest radiographs with RT-PCR–confirmed COVID-19 pneumonia from 2060 patients, and 5300 chest radiographs with non–COVID-19 pneumonia from 3148 patients (Figs 1, 2).

The chest radiography was performed on imaging systems from the following vendors: Carestream Health (DRX-1, DRX-Revolution), GE Healthcare (Optima-XR220, Geode Platform), Konica Minolta (CS-7), Agfa (DXD40, DXD30, DX-G), Siemens Healthineers (Fluorospot Compact FD), and Kodak (Classic CR).

Training, Validation, and Test Data Sets

It is important to consider any variables from chest radiograph acquisition (such as x-ray tube potential [ie, kilovoltpeak values] and x-ray exposure levels) to mitigate any biases in algorithm training (Appendix E1 [online]). Because our overarching objective was to develop a deep learning algorithm that could be successfully and broadly applied to chest radiographs obtained at different hospitals and clinics where chest radiography systems from different vendors are used, our strategy was to train the deep learning method by using a data set with images from different vendor systems. Chest radiographs were randomly selected from the four major vendors (Carestream Health, GE Healthcare, Konica Minolta, and Agfa) in the data set and these vendors were randomly anonymized as V1, V2, V3, and V4. The curated chest ra-

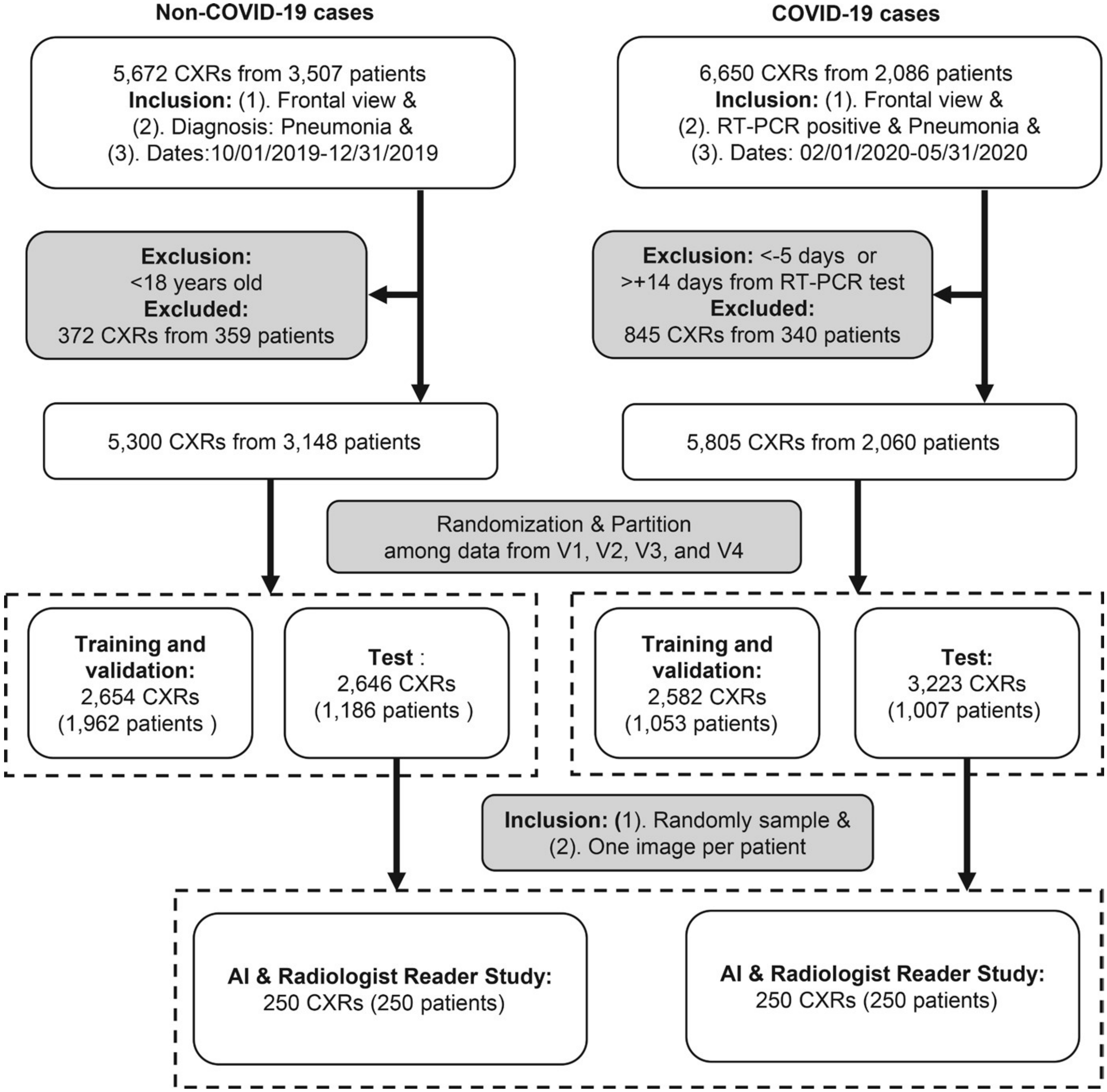

Figure 1: Study flowchart for data curation and data partition. Vendors 1–4 (V1–V4) are four major vendors of the acquired chest radiographs (CXRs) in the data set. AI = artificial intelligence, COVID-19 = coronavirus disease 2019, RT-PCR = reverse transcription polymerase chain reaction.

diographs were first grouped by vendors, and a total of 5236 chest radiographs (2582 chest radiographs in the COVID-19 cohort and 2654 chest radiographs in the non–COVID-19 pneumonia cohort) were used as training and validation to develop our deep learning algorithm, which is referred to as CV19-Net.

The remaining data were used for performance evaluation of the developed CV19-Net algorithm, including 3223 chest radiographs positive for COVID-19 from 1007 patients and 2646 non–COVID-19 pneumonia chest radiographs in 1186 patients. A patient-based data partition scheme was used to ensure that chest radiographs in any particular patient would only

Figure 2: Detailed data characteristics. **(a)** Age distribution of included patients. **(b)** Distribution of the Δ (delta; time between the positive reverse transcription polymerase chain reaction [RT-PCR] test and the chest radiography) for the positive cohort. A positive delta value indicates that the chest radiography was performed after the RT-PCR test. **(c)** Distribution of the radiographic unit vendors. **(d)** Distribution of the use of computed radiography (CR) or digital radiography (DX). **(e)** Distribution of data from different hospitals (H01–H05 indicate the five different hospitals and C01–C30 indicate the 30 different clinics). COVID-19 = coronavirus disease 2019.