Chest CT and Coronavirus Disease (COVID-19): A Critical Review of the Literature to Date

Constantine A. Raptis[1]
Mark M. Hammer[2]
Ryan G. Short[1]
Amar Shah[3]
Sanjeev Bhalla[1]
Andrew J. Bierhals[1]
Peter D. Filev[4]
Michael D. Hope[5,6]
Jean Jeudy[7]
Seth J. Kligerman[8]
Travis S. Henry[5]

Keywords: coronavirus, COVID-19, CT, infection, sensitivity, specificity

doi.org/10.2214/AJR.20.23202

Received March 25, 2020; accepted without revision March 26, 2020.

[1]Mallinckrodt Institute of Radiology, Washington University School of Medicine in Saint Louis, 510 S Kingshighway Blvd, Saint Louis, MO 63110. Address correspondence to C. A. Raptis (raptisc@wustl.edu).

[2]Department of Radiology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA.

[3]Department of Radiology, Zucker School of Medicine at Hofstra/Northwell, Manhasset, NY.

[4]Department of Radiology and Imaging Sciences, Emory University School of Medicine, Atlanta, GA.

[5]Department of Radiology and Biomedical Imaging, University of California, San Francisco, San Francisco, CA.

[6]Department of Radiology, San Francisco Veterans Affairs Medical Center, San Francisco, CA.

[7]Department of Diagnostic Radiology and Nuclear Medicine, University of Maryland School of Medicine, Baltimore, MD.

[8]Department of Radiology, University of California, San Diego, San Diego, CA.

AJR 2020; 215:1–4

ISSN-L 0361–803X/20/2154–1

OBJECTIVE. Coronavirus disease (COVID-19) is a global pandemic. Studies in the radiology literature have suggested that CT might be sufficiently sensitive and specific in diagnosing COVID-19 when used in lieu of a reverse transcription–polymerase chain reaction test; however, this suggestion runs counter to current society guidelines. The purpose of this article is to critically review some of the most frequently cited studies on the use of CT for detecting COVID-19.

CONCLUSION. To date, the radiology literature on COVID-19 has consisted of limited retrospective studies that do not substantiate the use of CT as a diagnostic test for COVID-19.

With increasing numbers of patients presenting with possible coronavirus disease (COVID-19), an efficient approach to triage is needed to conserve resources and mitigate the spread of disease. The role of CT as an adjunct to or replacement for reverse transcription–polymerase chain reaction (RT-PCR) in the screening or diagnosis of COVID-19 pneumonia has been the subject of much debate. The potential value of CT is that it is widely available and fast. RT-PCR, on the other hand, still is not readily available because of a lack of testing kits and reagents, and its turnaround times are variable, ranging from hours to days. The radiology literature has reported the characteristic CT findings of COVID-19 pneumonia, which most commonly include bilateral, peripheral, often-rounded ground-glass opacities that are predominantly located in the lower lobes and that may be accompanied by consolidation [1–12]. These reported findings of COVID-19 pneumonia are not unique or surprising; instead, they represent common but nonspecific imaging manifestations of acute lung injury with subsequent organization and are associated with numerous infectious and noninfectious inflammatory conditions [13, 14]. In a global pandemic, recognition of these CT findings is important in identifying patients with possible COVID-19 pneumonia who require further clinical evaluation, particularly when the findings are incidental and identified on CT scans for other indications.

The implications for the widespread deployment of CT in the evaluation and triage of patients with suspected COVID-19 pneumonia are less clear. The Centers for Disease Control and Prevention, the American College of Radiology, and the Society of Thoracic Radiology and American Society of Emergency Radiology have all issued position statements recommending against the use of CT for the screening and diagnosis of COVID-19 pneumonia, reserving CT for cases in which there is clinical suspicion of a complication of the disease or another diagnosis [15, 16]. These recommendations, however, are at odds with recently published studies reporting that CT has high sensitivity and specificity in the evaluation of suspected COVID-19 pneumonia [6, 11, 12].

The purpose of this article is to critically review some of the most frequently cited studies discussing the use of CT for COVID-19 and determine whether the current data justify a potential role for CT in the screening, diagnosis, or combined screening and diagnosis of COVID-19 and whether CT can function as a replacement for RT-PCR in areas where there is an outbreak of COVID-19 pneumonia or when the availability of RT-PCR testing is limited.

Sensitivity of CT in Detecting Coronavirus Disease Pneumonia

In a clinical scenario involving a communicable disease, the sensitivity of a diagnos-

tic test is important because misdiagnosis of even a single patient (i.e., obtaining a false-negative finding) can result in large outbreaks among future contacts. To our knowledge, the first study to report the sensitivity of CT in detecting COVID-19 pneumonia was a study by Fang et al. [6] that included 51 patients who were ultimately proven to have positive RT-PCR tests and had CT scans obtained at various time points during the course of their disease. Fang and colleagues retrospectively determined that 50 of the 51 patients (98%) had abnormal findings on baseline CT scans, whereas only 36 of the 51 patients (71%) had positive initial RT-PCR tests. The authors concluded that the "results support the use of chest CT for screening for COVID-19 for patients with clinical and epidemiologic features compatible with COVID-19 infection particularly when RT-PCR testing is negative" [6].

However, Fang and colleagues [6] did not fully discuss the limitations of their retrospective study, and the conclusion that they reached is unsubstantiated. One limitation of their study was selection bias. Patients in the cohort were selected to undergo CT on the basis of unknown clinical factors that may have distinguished them from similar patients who did not undergo CT. For instance, were the CT scans performed for patients in the imaging cohort because they had more severe symptoms? Did all patients with a suspected diagnosis of COVID-19 undergo CT or did only those with more severe symptoms do so? What about patients without symptoms? Given that the patients in the study received care at a hospital, it is possible and likely true that they had more severe symptoms. Unfortunately, the authors provide neither information regarding how the patients were selected to undergo CT nor data regarding patients who did not undergo CT.

The possibility that the cohort in the study by Fang et al. [6] is biased toward individuals with more severe disease is also suggested by a review of other studies in the literature. One such study, conducted by Inui et al. [8], reported that among patients on the Diamond Princess cruise ship who had COVID-19, CT findings were observed for 80% versus 46% of patients with and without symptoms, respectively [8]. In addition, Bernheim et al. [3] reported that normal CT findings were observed for 56%, 9%, and 4% of patients at 0–2, 3–5, and 6–12 days after symptom onset, respectively. After consideration of these studies and others, which report varying percentages of normal CT findings among patients with positive RT-PCR results [3, 4, 8, 10, 12], we consider the 2% rate of negative CT findings in the study by Fang and colleagues to be an outlier, suggesting that patient selection in that study was biased.

It should also be noted that a cohort with more severe symptoms would be different from a cohort without symptoms (i.e., the cohort typically used for a screening examination), or a cohort with risk factors or generalized symptoms worrisome for COVID-19. This is based on a broad description of the affected Chinese population in a study reporting that most patients with COVID-19 (81%) had mild symptoms [17]. Ultimately, understanding the potential effects of selection bias is important in determining sensitivity, because if a study cohort contains patients who are more likely to have a true-positive finding and less likely to have a false-negative finding, sensitivity will be overestimated.

A second shortcoming of the study by Fang et al. [6] was the lack of details regarding the definition of positive CT findings. The authors stated that 72% of patients had what were considered typical CT findings, whereas 28% had atypical findings. In reporting these atypical findings, the authors presented CT images (see Figs. 3A–3D in [6]) that showed focal ground-glass opacities as small as 5 mm [6]. These findings suggested that the threshold for a positive CT finding may have been abnormally low compared with the threshold used in standard practice; if so, this would have resulted in overestimation of the sensitivity of CT. In addition, the authors offered no evidence to suggest that the atypical findings were even related to COVID-19. In varying from standard practice by considering CT a binary test and using an abnormally low and nonclinical threshold, the authors overestimated the sensitivity of CT at the expense of specificity.

A study by Ai et al. [11], which presented a larger cohort of 1014 patients and reported a 97% sensitivity of CT in diagnosing COVID-19, had limitations similar to those in the study by Fang et al. [6]. The patient population in the study by Ai and colleagues was not clearly defined, but we inferred that the cohort included hospitalized patients who, compared with outpatients, are more likely to have abnormal CT findings [11]. As in the study by Fang and colleagues, CT was used as a binary test, and there was no defined threshold for determining positive examinations. The imaging examples provided by the authors showed focal ground-glass opacities that could be seen in many other diseases. Again, when such cases are considered to have positive CT findings, specificity is compromised and sensitivity is overestimated. The only major difference between the two studies was that the study by Ai and colleagues included a larger patient population, which would erroneously suggest to the reader that the cohort and study design were better equipped to evaluate sensitivity, when in fact the cohort and design used for this purpose were as limited as those used by Fang and colleagues. Both studies repeated the same errors and came to the same overreaching conclusions.

Both Fang et al. [6] and Ai et al. [3] also attempted to compare the sensitivity of CT to the sensitivity of RT-PCR in their cohorts. Fang and colleagues reported a time-to-positivity comparison and showed that of their 51 patients, 36 had initial positive RT-PCR findings, whereas 50 had initial positive CT findings; all 51 patients were reported to have positive RT-PCR test findings later in the course of the disease [6]. Ai and colleagues also conducted a time-to-positivity comparison and reported data on patients whose findings evolved from negative to positive RT-PCR test results: of 15 patients, 10 (67%) had positive findings on baseline CT scans [11]. These comparisons are difficult to interpret because of selection bias in the cohort and because CT was used as a binary test with an undefined and abnormally low threshold for positivity; both of these factors resulted in overestimation of the sensitivity of CT in detecting COVID-19.

Specificity of CT in Detecting Coronavirus Disease Pneumonia

We expect the reference standard test for any infection to be a laboratory test. The reason for this is straightforward: the specificity and positive predictive value of a test are based on its ability to limit false-positive findings. Laboratory tests (in this case, RT-PCR, which is a molecular assay) are able to test a feature of the disease that is not present in patients without the disease or in those who have other diseases. Although false-positive RT-PCR results are possible, they typically are caused by contamination and likely are negligible in the setting of assays for COVID-19. This is in contrast to CT, which

does not test for singular features unique to the disease. Even the features that are reported to be most characteristic of COVID-19 pneumonia (i.e., peripheral, bilateral ground-glass opacities that are predominantly found in the lower lobes) can be seen in a large number of other conditions, including other infectious and noninfectious conditions (as discussed later in this article).

With these considerations in mind, we believe that reports of CT having high specificity in diagnosing COVID-19 pneumonia should be viewed with skepticism. One such report is a retrospective study by Bai et al. [12], who focused on evaluating CT reader performance in distinguishing 219 cases of COVID-19 pneumonia at Chinese hospitals (37 cases were excluded because of normal CT findings) from 205 cases of viral pneumonia in the United States that had positive findings on respiratory viral panel tests. Bai and colleagues concluded that "radiologists are capable of distinguishing COVID-19 from viral pneumonia on chest CT with high specificity" [12]. The reported specificities (24%, 88%, and 94%) for CT readers who reviewed all cases varied greatly. The authors provided no explanation for the low specificity of the outlier radiologist, although that radiologist had a much higher sensitivity than the others and therefore was likely identifying most cases as COVID-19 regardless of imaging appearance.

If we ignore this outlier, radiologists who have experience reading thoracic CT examinations may be surprised at the high specificities (87% and 92%) achieved in the diagnosis of any infectious pneumonia. A closer examination of the study design and patient cohorts in the study by Bai et al. [12] offers reasons for these results. First, the control group included only patients with viral pneumonias. No patients had noninfectious diseases with findings that might have overlapped with CT findings of COVID-19 pneumonia, such as pulmonary edema, organizing pneumonia, or lung injury of any other cause (e.g., drug toxicity, radiation treatment, or a cryptogenic cause), pulmonary infarcts, alveolar hemorrhage, and interstitial lung diseases (e.g., nonspecific interstitial pneumonia or desquamative interstitial pneumonia). If the control group included no patients with diseases other than viral pneumonias that may have overlapped with COVID-19 pneumonia, specificity would be overestimated.

Second, no objective criteria were used to define a positive CT examination, and the diagnosis was reached by gestalt, as was reflected by the very low performance of one of the readers.

Finally, the study did not describe any training CT readers may have received before reviewing the images. In real-world practice, the differential diagnosis for a patient presenting with fever and cough is quite broad, and other aforementioned causes are a significant proportion of the diagnoses. Consequently, the specificity of chest CT for the diagnosis of COVID-19 is almost certainly much lower than that reported by Bai and colleagues [12], and the lack of any training information or specific diagnostic criteria in their study limits the application of CT in real-world clinical practice.

Implications for the Application of CT in Clinical Practice

We have addressed the limitations of some of the most frequently cited studies in the radiology literature about the use of CT for evaluating patients suspected of having COVID-19 pneumonia. In our opinion, the studies to date provide no compelling data to support the use of CT as a screening test for populations with symptoms or for those suspected of having the disease. The sensitivity of CT varies widely, and none of the studies adequately evaluate the use of CT in a representative screening population [3, 4, 10, 12]. The studies reporting high sensitivity values are fraught with selection bias, and they vary from clinical practice in that they consider CT a binary test with abnormally low thresholds for positive results. These factors lead to an overestimation of the sensitivity of CT in the diagnosis of COVID-19 pneumonia. As such, the negative CT results cannot be reliably believed. As previously stated, the consensus of many national and international organizations, including the Centers for Disease Control and Prevention, the American College of Radiology, and the Society of Thoracic Radiology and American Society of Emergency Radiology, has also affirmed this conclusion.

None of the literature reviewed in this article reliably reports a high specificity of CT in differentiating COVID-19 pneumonia from other diseases with similar CT findings, thereby limiting the use of CT as a confirmatory diagnostic test. In populations of patients with a high prevalence of COVID-19 pneumonia (as in disease surges or outbreaks), the positive predictive value of CT will appear increased even if CT is not specific. However, according to the Bayes theorem, positive CT results are unlikely to be clinically useful because the posttest probability would not be significantly different from the high pretest probability, given the overall high prevalence of disease.

Even in situations in which RT-PCR test results are negative, delayed, or not available, no data of which we are aware support CT as an adequate replacement test because its true sensitivity is unknown (and is unlikely to be of value given the known existence of normal CT findings in patients with the disease) and because CT findings lack specificity. In other words, no high-quality data of which we are aware support the wide deployment of CT to meaningfully improve the management of patients with suspected COVID-19 pneumonia.

Finally, no diagnostic test is without risks, and the hazards of wide deployment of CT must be acknowledged. These risks include overuse of hospital resources, including the use of protective gear that is already limited in availability but is required to safely perform CT studies; clustering of affected and nonaffected patients in imaging departments, thereby potentially increasing risks of disease transmission and exposure among staff performing the examinations; patients for other indications because of increased use of CT scanners to evaluate cases of suspected COVID-19.

Conclusion

To date, the studies reporting CT features of COVID-19 pneumonia have been retrospective reviews and case series. They should be considered low quality, providing a level 3 body of evidence [18]. This is not to say these studies are not valuable. Reports of the various CT features of COVID-19 pneumonia are an important first step in helping radiologists identify patients who may have COVID-19 pneumonia in the appropriate clinical environment. However, test performance and management issues arise when inappropriate and potentially overreaching conclusions regarding the diagnostic performance of CT for COVID-19 pneumonia are based on low-quality studies with biased cohorts, confounding variables, and faulty design characteristics.

At present, CT should be reserved for evaluation of complications of COVID-19 pneumonia or for assessment if alternative diagnoses are suspected. As the medical community gains experience in treating patients with COVID-19 pneumonia, high-

quality data hopefully will emerge and will support a more expanded role for CT. We (and the radiology community at large) will welcome any such data to improve the care of patients with this disease.

References

1. Shi H, Han X, Jiang N, et al. Radiological findings from 81 patients with COVID-19 pneumonia in Wuhan, China: a descriptive study. *Lancet Infect Dis* 2020; 20:425–434
2. Zhou S, Wang Y, Zhu T, Xia L. CT features of coronavirus disease 2019 (COVID-19) pneumonia in 62 patients in Wuhan, China. *AJR* 2020 Mar 5 [Epub ahead of print]
3. Bernheim A, Mei X, Huang M, et al. Chest CT findings in coronavirus disease-19 (COVID-19): relationship to duration of infection. *Radiology* 2020 Feb 20 [Epub ahead of print]
4. Chung M, Bernheim A, Mei X, et al. CT imaging features of 2019 novel coronavirus (2019-nCoV). *Radiology* 2020; 295:202–207
5. Fang Y, Zhang H, Xu Y, Xie J, Pang P, Ji W. CT manifestations of two cases of 2019 novel coronavirus (2019-nCoV) pneumonia. *Radiology* 2020; 295:208–209
6. Fang Y, Zhang H, Xie J, et al. Sensitivity of chest CT for COVID-19: comparison to RT-PCR. *Radiology* 2020 Feb 19 [Epub ahead of print]
7. Han R, Huang L, Jiang H, Dong J, Peng H, Zhang D. Early clinical and CT manifestations of coronavirus disease 2019 (COVID-19) pneumonia. *AJR* 2020 Mar 17 [Epub ahead of print]
8. Inui S, Fujikawa A, Jitsu M, et al. Chest CT findings in cases from the cruise ship "Diamond Princess" with coronavirus disease 2019 (COVID-19). *Radiol Cardiothorac Imaging* 2020 Mar 17 [Epub ahead of print]
9. Liu KC, Xu P, Lv WF, et al. CT manifestations of coronavirus disease-2019: a retrospective analysis of 73 cases by disease severity. *Eur J Radiol* 2020; 126:108941
10. Pan F, Ye T, Sun P, et al. Time course of lung changes on chest CT during recovery from 2019 novel coronavirus (COVID-19) pneumonia. *Radiology* 2020 Feb 13 [Epub ahead of print]
11. Ai T, Yang Z, Hou H, 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 Feb 13 [Epub ahead of print]
12. Bai HX, Hsieh B, Xiong Z, et al. Performance of radiologists in differentiating COVID-19 from viral pneumonia on chest CT. *Radiology* 2020 Mar 10 [Epub ahead of print]
13. Franquet T. Imaging of pulmonary viral pneumonia. *Radiology* 2011; 260:18–39
14. Kligerman SJ, Franks TJ, Galvin JR. From the radiologic pathology archives: organization and fibrosis as a response to lung injury in diffuse alveolar damage, organizing pneumonia, and acute fibrinous and organizing pneumonia. *RadioGraphics* 2013; 33:1951–1975
15. American College of Radiology (ACR). ACR recommendations for the use of chest radiography and computed tomography (CT) for suspected COVID-19 infection. ACR website. www.acr.org/Advocacy-and-Economics/ACR-Position-Statements/Recommendations-for-Chest-Radiography-and-CT-for-Suspected-COVID19-Infection. Updated March 2020. Accessed March 24, 2020
16. American Society of Emergency Radiology (ASER). ASER COVID-19 Task Force: FAQs. ASER website. www.aser.org/covid-19-faqs/. Published March 2020. Accessed March 24, 2020
17. Wu Z, McGoogan JM. Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China: summary of a report of 72 314 cases from the Chinese Center for Disease Control and Prevention. *JAMA* Network website. jamanetwork.com/journals/jama/fullarticle/2762130. Published February 24, 2020. Accessed March 23, 2020
18. American College of Radiology (ACR). ACR Appropriateness Criteria. ACR website. www.acr.org/Clinical-Resources/ACR-Appropriateness-Criteria. Updated March 2020. Accessed March 24, 2020