Determinants of Chest X-Ray Sensitivity for COVID- 19: A Multi-Institutional Study in the United States

Stephanie Stephanie, MD[1*], Thomas Shum, MD, PhD[1*], Heather Cleveland[2], Suryanarayana R. Challa, MD[1], Allison Herring, MD[3], Francine L Jacobson MD, MPH[5], Hiroto Hatabu MD, PhD[5], Suzanne C Byrne, MD[5], Kumar Shashi, MB, BS[5], Tetsuro Araki, MD, PhD[5], Jose A. Hernandez, MD[4], Charles S. White, MD[3], Rydhwana Hossain, MD[3], Andetta R Hunsaker, MD[5], Mark M Hammer, MD[5]

1. Department of Internal Medicine, University of Maryland School of Medicine, Midtown Campus, 827 Linden Avenue, Baltimore, MD 21201.

2. Department of Physician Assistant Studies, Massachusetts General Hospital Institute of Health Professions, 55 Fruit St, Boston, MA 02114.

3. Department of Radiology, University of Maryland School of Medicine, Downtown Campus, 22 S Greene St, Baltimore, MD 21201.

4. Department of Pediatric Radiology, Texas Children's Hospital, 6621 Fannin St, Houston, TX 77030.

5. Department of Radiology, The Brigham and Women's Hospital, Harvard Medical School, 75 Francis St, Boston, MA 02114.

* These co-authors contributed equally to the publication

Corresponding Author: Mark Hammer, MD

Department of Radiology

The Brigham and Women's Hospital, Harvard Medical School

75 Francis St, Boston MA 02115

Email: mmhammer@bwh.harvard.edu

MANUSCRIPT TYPE: Original Research

KEY RESULTS

- Sensitivity of CXR for COVID-19 increases from 55% at ≤2 days to 79% at >11 days after symptom onset.

- Accuracy of diagnosis of COVID-19 in serial CXR approaches that of chest CT.

- Normal or mild severity CXRs were the main determinants of false negative CXR interpretations.

SUMMARY STATEMENT

Serial chest radiography sensitivity for detection of novel 2019 coronavirus infection approaches chest CT, suggesting utility of chest radiography as an adjunctive diagnostic tool for COVID-19 patients.

ABBREVIATIONS

AUC (area under the curve), ROC (receiver operating characteristic curve), COVID-19 (novel coronavirus 2019 infection), COVID-19+ (positive for novel coronavirus 2019 infection), COVID-19- (negative for novel coronavirus 2019 infection), GGO (ground-glass opacity), DA (diffuse airspace opacity), IA (interstitial and airspace opacities), FN (false negative)

Abstract

Purpose: To evaluate the sensitivity, specificity, and severity of chest x-rays (CXR) and chest CTs over time in confirmed COVID-19+ and COVID-19- patients and to evaluate determinants of false negatives.

Methods: In a retrospective multi-institutional study, 254 RT-PCR verified COVID-19+ patients with at least one CXR or chest CT were compared with 254 age- and gender-matched COVID-19- controls. CXR severity, sensitivity, and specificity were determined with respect to time after onset of symptoms; sensitivity and specificity for chest CTs without time stratification. Performance of serial CXRs against CTs was determined by comparing area under the receiver operating characteristic curves (AUC). A multivariable logistic regression analysis was performed to assess factors related to false negative CXR.

Results: COVID-19+ CXR severity and sensitivity increased with time (from sensitivity of 55% at ≤ 2 days to 79% at >11 days; $p<0.001$ for trends of both severity and sensitivity) whereas CXR specificity decreased over time (from 83% to 70%, $p=0.02$). Serial CXR demonstrated increase in AUC (first CXR AUC=0.79, second CXR=0.87, $p=0.02$), and second CXR approached the accuracy of CT (AUC=0.92, $p=0.11$). COVID-19 sensitivity of first CXR, second CXR, and CT was 73%, 83%, and 88%, whereas specificity was 80%, 73%, and 77%, respectively. Normal and mild severity CXR findings were the largest factor behind false-negative CXRs (40% normal and 87% combined normal/mild). Young age and African-American ethnicity increased false negative rates.

Conclusion: CXR sensitivity in COVID-19 detection increases with time, and serial CXRs of COVID-19+ patients has accuracy approaching that of chest CT.

Introduction

In December 2019, the first outbreak of the novel 2019 coronavirus disease (COVID-19) was reported in China in Wuhan City, Hubei Province(1). It rapidly achieved pandemic status, and the United States currently has the highest number of cases in the world. In the US, the medical community has been primarily focused on controlling viral spread through mitigation measures(2). Until favorable developments occur from drug discovery and vaccine development efforts, rapid and sensitive detection of COVID-19 infection is essential for triage of patients, both for timely isolation of COVID-19+ patients and for appropriate allocation of hospital resources and providers in hospital units with limited capacity.

The gold standard for diagnostic confirmation of COVID-19 infection in patients is a positive result from reverse transcriptase polymerase chain reaction (RT-PCR), but the time-to-result from labs (which may be as long as several days) has been a rate-limiting step in hospital care. As a result, imaging-based diagnostic tools are attractive as a companion option, with some reports suggesting high sensitivity of COVID-19 detection by computed tomography (CT) studies(3,4). However, questions have been raised about the reported high sensitivity of CT in those studies, which likely suffer from selection bias toward more severe cases(5). Some early reports have also described high specificity in diagnosis of COVID-19 compared to other entities, but these too may be biased by selection of more severe cases and, at least in one case, an artificial control group of other viral pneumonias(6,7). In addition, the ability to detect COVID-19 pneumonia on imaging likely varies over time but remains to be fully elucidated(8).

Here, we evaluate the imaging characteristics, sensitivity, and specificity of both chest X-ray (CXR) and chest CT in detecting COVID-19 infection in patients with lab-confirmed SARS-CoV-2 viral infection over time. We also evaluate the contributions of patient demographics and comorbidities that influenced false-negative imaging.

Methods

Patient selection. This retrospective study was conducted at two large urban academic medical centers, including three tertiary care hospitals and one community hospital. The respective Institutional Review Boards at the two academic medical centers approved the study, with waiver of informed consent. Inclusion criteria were patients who had both at least 1 nucleic acid amplification-based COVID-19 detection test and at least 1 CXR or CT that was performed within a week of testing. Patients younger than 18 years of age were excluded. A patient was regarded as COVID-19 positive if they had any positive test within the time interval. At the first institution, 182 patients positive for COVID-19 infection from March 1, 2020 through April 1, 2020 were randomly selected and then paired with age- and gender-matched control patients who tested COVID-19 negative. At the second institution, 72 positive for COVID-19 infection between March 1, 2020 through April 14, 2020 and age- and gender-matched control patients were then included in a similar process.

Image review

At each site, two board-certified, fellowship-trained thoracic radiologists reviewed each study separately. At the first institution, images were reviewed by two of the thoracic radiologists (MMH, ARH, FLJ, HH, and SCB who had between 5 and 29 years of experience). At the second institution, RAH with 4 years of experience and CW with 29 years of experience reviewed the studies. Reviewers were blinded to the diagnosis.

Chest X-Ray evaluation. Reviewers evaluated for the predominant pattern in chest radiographs, selecting from: interstitial opacities, interstitial and airspace opacities, atelectasis, diffuse airspace opacities, lobar consolidation, or peripheral opacities; if none of these patterns applied the reviewers would select 'normal'. Reviewers also assessed for a craniocaudal gradient and the presence of pleural effusions. Reviewers were asked to assign a severity score for the chest x-ray findings: normal, mild, moderate, or

severe. Reviewers also assessed the likelihood of COVID-19 based on the radiograph findings, with a score between 1 and 5 (1 being "Very unlikely," 3 being "Intermediate likelihood," and 5 being "Highly Likely"). Up to 3 serial chest radiographs were scored per patient.

Chest-CT evaluation. Reviewers were asked to look for the presence of nodules, ground glass opacities, consolidation, and septal thickening. The presence of pleural effusions and reverse halo sign were also assessed. Additionally, reviewers evaluated craniocaudal gradient, anterior/posterior gradient, and central/peripheral gradient; reviewers could score cases as having a gradient or diffuse involvement. The reviewers were asked to assign a chest CT severity score and a chest CT COVID-19 likelihood score.

Patient Demographics. Patient's gender, age, ethnicity, height, and admission weight were collected. Patient history was reviewed for positive history for hypertension, diabetes mellitus, cardiac disease or anomaly, chronic lung disease, autoimmune disease, malignancy, stem cell transplant, solid organ transplant, or chronic kidney disease.

Data Entry and Statistics.

Data was entered using the Castor EDC platform(9) and the REDCap platform(10). Graphpad Prism 5.0 and Graphpad Prism 8.4.2 (Graphpad Software, San Diego, CA) as well as JMP (v15, SAS Institute, Cary, NC) statistical software packages were used for formal analysis of data.

In order to combine data from the two reviewers, for craniocaudal gradient, anterior/posterior gradient, and central/peripheral gradient, diffuse distribution was assigned to a case if 1) at least 1 reviewer scored diffuse or 2) if both reviewers selected opposing categories (for example, if Reviewer 1 selected anterior distribution and Reviewer 2 selected posterior distribution). CXRs and chest CTs were characterized as having normal patterns only if both reviewers characterized it as normal. If the two reviewers selected different abnormal patterns, both patterns were scored for purposes of analysis; if one reviewer selected

an abnormal pattern while the other selected 'normal', then only the abnormal pattern was scored. For severity and COVID-19 likelihood scores, the scores were averaged for the two reviewers.

For purposes of sensitivity and specificity, a CXR or CT scored as 3 (intermediate COVID-19 likelihood) or above in the COVID-19 likelihood score was classified as a 'positive' test.

The time intervals between symptom onset and chest radiographs were evaluated, and the quartiles were used to generate time interval bins for analysis of CXR and CT changes over time. Comparisons of categorical variables were done with Fisher's exact test, one-tailed, for 2x2 tables or the chi-squared test for variables with multiple levels. Trend analysis of categorical variables across time interval bins was done using the Cochran-Armitage Test. Comparison of continuous variables across time interval bins was done with ANOVA. Multivariable logistic regression analysis was performed to evaluate factors associated with false negative chest x-rays. A p-value threshold of 0.05 was used for statistical significance. To adjust for multiple comparisons, we applied the Benjamini-Hochberg correction to the p-values for each table of univariable analyses, using a false-discovery rate of 0.05, and we named the p-values "Adjusted p-value" in these instances.

Confidence intervals and standard errors of the areas under the receiver operating characteristic curves (AUC) were obtained using bootstrapping with 2500 samples. The comparison of paired AUCs was performed using a previously described method(11).

Results

Patient characteristics. A total of 508 patients were included, 254 of whom were COVID-19+ with 254 matched COVID-19- controls. The demographics and clinical characteristics of the cohort are given in Table 1. 500 (98%) of patients had at least 1 CXR, 195 (38%) had at least 2 CXRs, 112 (22%) had 3 CXRs, and 169 (33%) had at least one chest CT. CXR were obtained at the following time intervals after

onset of symptoms: 254 (32%) at ≤2 days, 192 (24%) at 3-6 days, 183 (23%) at 7-11 days, and 177 (22%) at >11 days. The median time from symptoms to CT was 4 days. The available clinical indications for CXR and CT between COVID-19 positive and negative patients are listed in Supplemental Tables 1-4.

CXR evolution over time. The severity of CXR scores increased over time for COVID-19+ patients (from mean of 1.11 at ≤2 days to mean of 1.97 at >11 days, p<0.001, Figure 1A). However, the severity scores for COVID-19- patients did not change significantly over time (from mean of 0.93 at ≤2 days to mean of 1.10 at >11 days, p=0.21.

The most common CXR finding in COVID-19+ patients was interstitial and airspace opacities (IA pattern), present in 62% of cases. IA pattern increased over time, present in 51% at ≤2 days versus 73% at >11 days (p<0.001 for trend, Figure 1B). CXR disease distribution also changed over time, with increasing prevalence of diffuse opacities (3% at ≤2 days vs 25% at >11 days, p<0.001 for trend, Figure 1C). Examples of evolving CXR in COVID-19+ patients are given in Figure 2.

Sensitivity of imaging over time. The rise of CXR severity over the course of COVID-19 infection was reflected in the sensitivity CXR COVID-19 detection over time (Figure 3A). The sensitivity of CXR at ≤2 days is 55%, increasing to 79% at >11 days (p<0.001 for trend). Specificity declined slightly over time, 83% at ≤2 days to 70% at >11 days (p=0.02 for trend).

We evaluated sequential imaging tests per patient for diagnosis of COVID-19. The first chest radiographs (CXR1) were obtained a median of 3 days after symptoms, the second chest radiographs (CXR2) were obtained a median of 8 days after symptoms, and CTs were obtained a median of 4 days after symptoms. In patients with both CXR1 and CXR2 (n=195), areas under the receiver operating characteristic curve (AUC) were 0.80 for CXR1 and 0.871 for CXR2. In patients who had both CXR1 and a CT study performed (n=160), CXR1 AUC was 0.80 and CT AUC was 0.90. In patients who had CXR1, CXR2, and

CT (n=85), AUC was 0.79 (95% CI 0.67 – 0.88) for CXR1, 0.88 (95% CI 0.79 – 0.93) for CXR2, and 0.92 (95% CI 0.84 – 0.97) for CT (Figure 3B). CXR2 and CT AUCs were both significantly higher than the CXR1 AUC (p=0.02 and p=0.003, respectively), whereas comparison of CXR2 and CT AUCs found no significant difference (p=0.11). In this group, CXR1 had a sensitivity of 73% and specificity of 80%; CXR2 had sensitivity of 83% and specificity of 73%; and CT had sensitivity of 88% and specificity of 77%.

Specific CXR1 findings for COVID-19. Frequencies of CXR1 findings in COVID-19 cases and controls are shown in Table 2 (with frequencies from total CXR shown in Supplemental Table 5). As noted above the most common CXR1 finding in COVID-19 was the IA pattern, present in 51% of COVID-19+ cases, versus 26% of COVID-19- cases (p<0.001). Peripheral airspace opacities and diffuse airspace disease were also more common in COVID-19+ cases than controls (21% versus 4%, p<0.001, and 6% versus 2%, p = 0.03). Conversely, COVID-19+ CXR1s demonstrated a reduced frequency of atelectasis, lobar consolidation, and pleural effusions compared to COVID-19- CXR1s (p=<0.001, p=0.03, p<0.001, respectively).

Specific CT findings for COVID-19. Frequencies of single CT findings in COVID-19+ cases and controls are given in Table 3. Diffuse involvement across craniocaudal and anterior/posterior spatial distributions of chest-CT findings predominated in COVID-19+ patients relative to COVID-19- patients (58% versus 28%, p<0.001, and 69% versus 47%, p=0.01, respectively), as did ground glass opacities (88% versus 49%, p<0.001). Additionally, the reverse halo sign was significantly more prevalent in COVID-19+ patients (33% in COVID-19+ cases versus 5% COVID-19- cases, p<0.001). Among these COVID-19+ characteristics in total CTs, only the frequency of consolidation significantly changed over the course of COVID-19 infection (Figure 4), increasing from 32% at ≤ 2 days to 91% at >11 days (p<0.001 for trend).

Features associated with false negative imaging. There were 113 false negative (FN) initial CXRs and 23 FN second CXRs in COVID-19+ patients. Normal CXRs comprised 51 (45%), 3 (14%), and 54 (40%) of FN initial CXRs, second CXRs, and the summed initial and second CXRs. The combined normal and mild severity FN CXRs for these three respective groups were 106 (94%), 9 (39%), and 118 (88%). Of the 113 patients with FN CXR1 studies, 35 had a CXR2, of which 16/35 were positive. Analysis of patient demographics and comorbidities initially identified female sex, chronic lung disease, absence of hypertension, ethnicity, and young age to be associated with FN initial CXR (Table 4). A multivariable logistic regression analysis showed young age and African-American ethnicity increased FN rates whereas Hispanic-American ethnicity decreased FN rates.

Of 10 FN CTs, 2 were normal and 7 (70%) were either normal or mild severity. Female sex (OR 5.44, p=0.03) and chronic lung disease (OR 5.56, p=0.02) also contributed to FN chest-CT studies.

Discussion

In summary, we evaluated the imaging characteristics, evolution of severity, sensitivity, and specificity of Chest X-Rays (CXRs) and chest CT in lab-verified novel 2019 coronavirus (COVID-19) positive and matched negative patients from two academic medical centers in two major US cities. Our results indicate that the sensitivity of CXR imaging increases over the course of COVID-19 infection, from as low as 55% at ≤2 days to 79% at >11 days after symptom onset, and serial CXR imaging approaches the diagnostic accuracy of chest CT (AUC of 0.875 versus 0.916). Normal imaging was the most prevalent factor underlying false negative (FN) CXRs and chest CTs. Patient characteristics including young age and certain ethnicities were associated with greater rates of FN exams. Finally, we found a number of specific imaging findings in CXR and CT that were helpful in the diagnosis of COVID-19, including interstitial and airspace opacity pattern (OR 2.84) and peripheral airspace opacities (OR 7.94) on CXR, and reverse halo sign on chest-CT (OR 9.36).

We showed that the severity and sensitivity of COVID-19+ CXR findings increases over time, with sensitivity markedly increasing after Day 6 following symptom onset. This coincides with the known time course of pathophysiologic deterioration in patients, as the time from illness onset to dyspnea, sepsis, and acute respiratory distress syndrome in COVID-19+ patients has been reported as 7 days, 9 days, and 12 days, respectively(12). A previous study analyzed baseline CXRs from 64 COVID-19+ patients and found a sensitivity of 44/64 (69%) in COVID-19 detection, comparable to our results(13). In our study, we stratified CXRs based on time since the onset of symptoms and were able to identify when imaging is expected to be most useful. Due to inadequate number of CTs in our patient population, we could not perform a similar analysis for CT sensitivity.

From our study population, we determined that the overall sensitivity of chest-CT for COVID-19 infection was 88%, which is comparable to 3 other studies reporting 86%, 86%, and 88% but lower than

studies reporting sensitivities as high as 100%(1,3,4,14–16). Of note, our article specifically addressed diagnostic accuracy within a relevant clinical scenario (distinguishing COVID-19+ from COVID-19- patients), rather than assessing sensitivity alone without a control group or with artificially generated control groups.

Our finding that serial CXR imaging approaches the accuracy of CT imaging in COVID-19 has potential value to enhancing bio-containment and workflow in hospitals. The American College of Radiology has recommended decontamination of CT-imaging rooms after each use for COVID-19 evaluation to prevent viral transmission from CT scanners(17). In addition to the significant cleaning duties and personal protective equipment donning/doffing required of radiology technicians and janitorial staff, the process of transporting inpatients to a commonly utilized CT unit accumulates risk for viral transmission to other patients and healthcare staff. These concerns potentially limit widespread adoption of CT imaging for COVID-19 detection. Conversely, CXRs uses less radiation and portable CXR machines can be brought to patient bedsides to perform CXRs and require far less time to decontaminate, making it an attractive option to employ.

The majority of false negative studies were caused by CXRs or CTs characterized as normal, which likely results from imaging taken during early viral infection when infectious changes are not yet visible in lung parenchyma. We observed that older COVID-19+ patients had a higher frequency of true positive CXRs, which corresponds to reports that COVID-19 has greater severity in older patients(12,18). We noted increased FN rates in African Americans and a decreased FN rate in Hispanic-Americans, relative to Caucasian-Americans. This may be related to differential testing in these minority populations, with some patients having access to testing for more mild cases and others not; alternatively, there may be true biological differences (or different rates of comorbidities) that contribute to different CXR finding severities. Regardless, these findings may have implications for equity of care across minority populations.

Our study has several limitations, including its retrospective nature, as there was not a consistent clinical or imaging algorithm for these patients. Other limitations lie in patient selection, specifically in that the reasons for patient imaging were unknown and that serial chest X-ray imaging and CT imaging is likely biased towards patients with more severe disease. However, the ROC comparisons performed in this study were done among patients who had the same set and number of imaging tests, so those patients should be comparable to each other.

In conclusion, the results from our study imply that strategic serial CXR imaging can detect COVID-19 infection in symptomatic patients and may be a helpful adjunct to RT-PCR testing. In an inpatient hospital setting in which testing kit availability, laboratory resources, and/or laboratory staff are compromised and threaten delay to patient care and hospital workflow, serial CXRs could potentially be utilized in an adjunct diagnostic role in patients with suspected COVID-19.

References

1. Huang C, Wang Y, Li X, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet. 2020;395(10223):497–506.

2. Sanders JM, Monogue ML, Jodlowski TZ, Cutrell JB. Pharmacologic Treatments for Coronavirus Disease 2019 (COVID-19): A Review. JAMA - J Am Med Assoc. 2020;323(18):1824–1836.

3. 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;2019:200642.

4. Fang Y, Zhang H, Xie J, et al. Sensitivity of Chest CT for COVID-19: Comparison to RT-PCR. Radiology. 2020;200432.

5. Raptis CA, Hammer MM, Short RG, et al. Chest CT and Coronavirus Disease (COVID-19): A Critical Review of the Literature to Date. AJR. 2020;215:1–4.

6. Bai HX, Hsieh B, Xiong Z, et al. Performance of radiologists in differentiating COVID-19 from viral pneumonia on chest CT. Radiology. 2020;200823.

7. Prokop M, van Everdingen W, van Rees Vellinga T, et al. CO-RADS - A categorical CT assessment scheme for patients with suspected COVID-19: definition and evaluation. Radiology. 2020;(1):201473.

8. Bernheim A, Mei X, Huang M, et al. Chest CT Findings in Coronavirus Disease-19 (COVID-19): Relationship to Duration of infection. Radiology. 2020;200463.

9. Castor EDC. (2019). Castor Electronic Data Capture. [online] Available at: https://castoredc.com. .

10. Harris P, Taylor R, Thielke R, Payne J, Gonzalez N, Conde J. Research electronic data capture (REDCap) – A metadata-driven methodology and workflow process for providing translational research informatics support,. J Biomed Inform. 2009;42(2):377–381.

11. Hanley JA, McNeil BJ. A method of comparing the areas under receiver operating characteristic

curves derived from the same cases. Radiology. 1983;148(3):839–843.

12. Zhou F, Yu T, Du R, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. Lancet. 2020;395(10229):1054–1062.

13. Wong HYF, Lam HYS, Fong AH-T, et al. Frequency and Distribution of Chest Radiographic Findings in COVID-19 Positive Patients Authors. Radiology. 2020;201160.

14. Chen N, Zhou M, Dong X, et al. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. Lancet. 2020;395(10223):507–513.

15. Song F, Shi N, Shan F, et al. Emerging 2019 novel coronavirus (2019-NCoV) pneumonia. Radiology. 2020;295(1):210–217; https://doi.org/10.1148/radiol.2020200274.

16. Chung M, Bernheim A, Mei X, et al. CT Imaging features of 2019 novel coronavirus pneumonia (2019-nCoV). Radiology. 2020;295:202–207; https://doi.org/10.1148/radiol.2020200230.

17. Jacobi A, Chung M, Bernheim A, Eber C. Portable chest X-ray in coronavirus disease-19 (COVID-19): A pictorial review. Clin Imaging. 2020;64(April):35–42.

18. Liu K, Chen Y, Lin R, Han K. Clinical features of COVID-19 in elderly patients: A comparison with young and middle-aged patients. J Infect. 2020;https://doi.org/10.1016/j.jinf.2020.03.005.
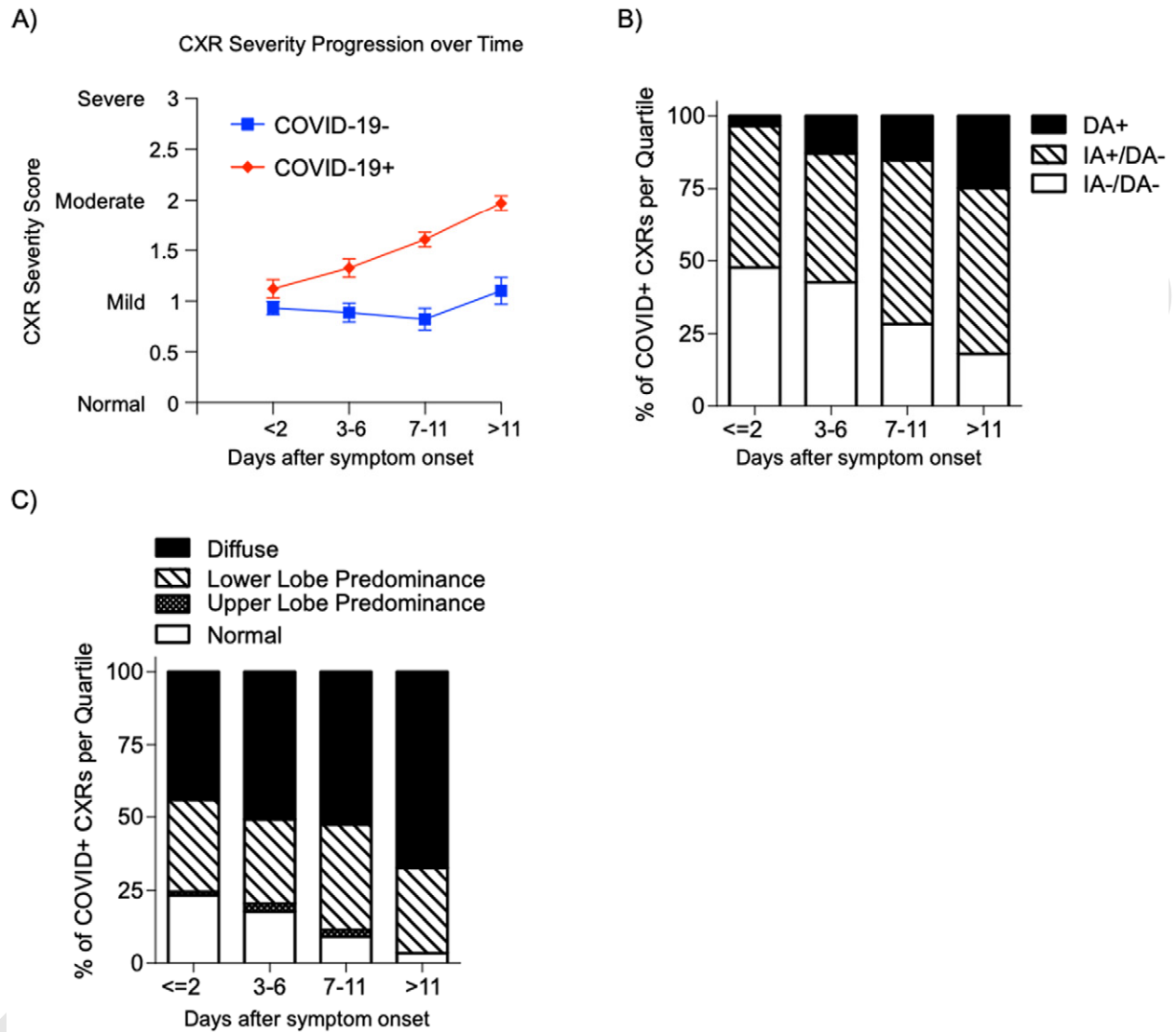
Figure 1

Figure 1. Evolution of COVID-19 chest X-ray severity, findings, and distribution after onset of symptoms. **(a)** Chest X-ray severity scores in COVID-19 infected patients (red) and COVID-19 negative (blue) patients after onset of symptoms. Graph represents averages from patients in each time interval ± standard error. **(b)** Frequencies of COVID-19+ chest X-rays positive for diffuse airspace opacities (DA) pattern, interstitial and airspace opacities pattern (IA) alone, or neither, in each time interval. **(c)** Frequencies of COVID-19+ chest X-rays positive for diffuse parenchymal involvement, lower lobe

predominance, upper lobe predominance, or distribution characteristic of a normal chest radiograph, in each time interval.
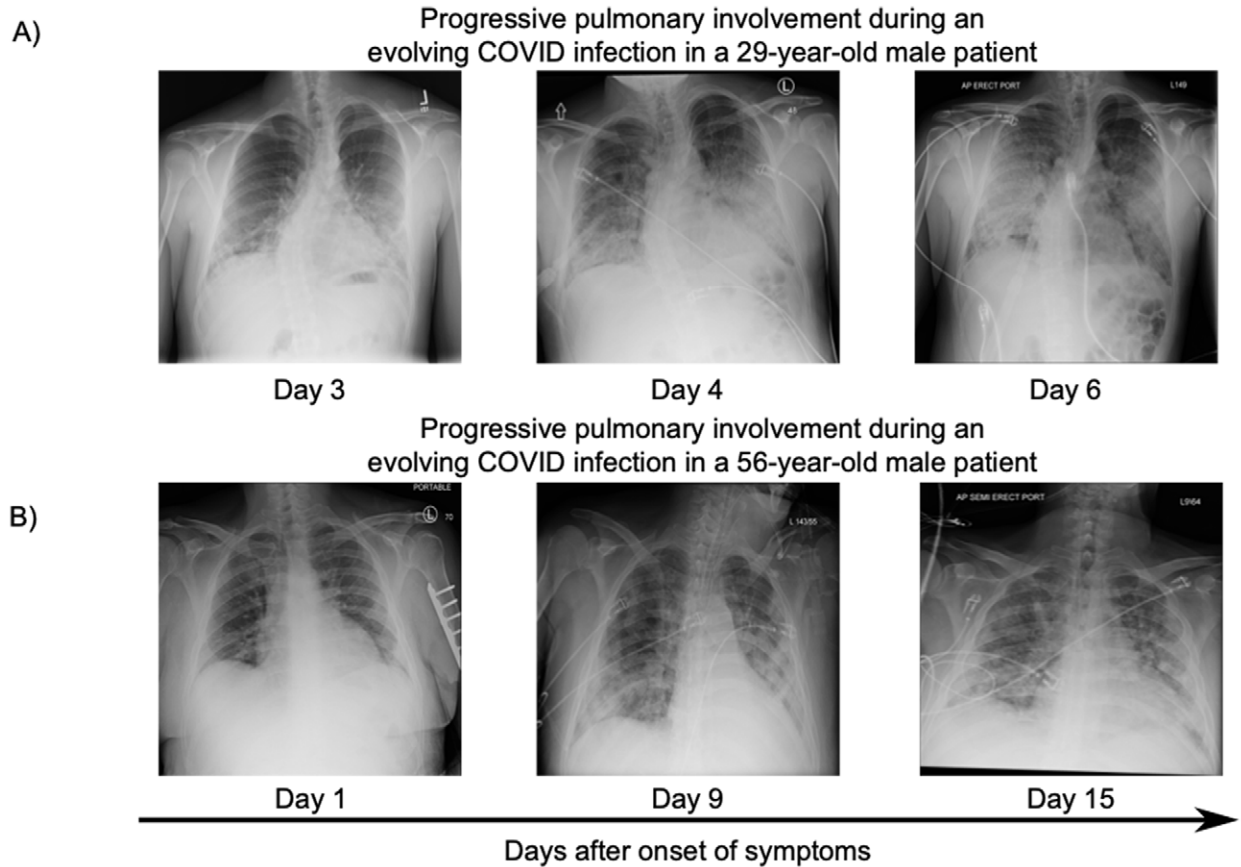

Figure 2. Representative serial chest X-ray imaging of COVID-19 infected patients. (a) 29-year-old man with rapid respiratory deterioration after symptom onset showing progression from lower lung predominant interstitial and airspace opacities on day 3 to diffuse involvement with extensive airspace disease on days 4 and 6. (b) 56-year-old-man with COVID-19, presenting initially with normal CXR, which then progressed to lower lung predominant interstitial and airspace opacities at day 9, which mildly worsened by day 15.
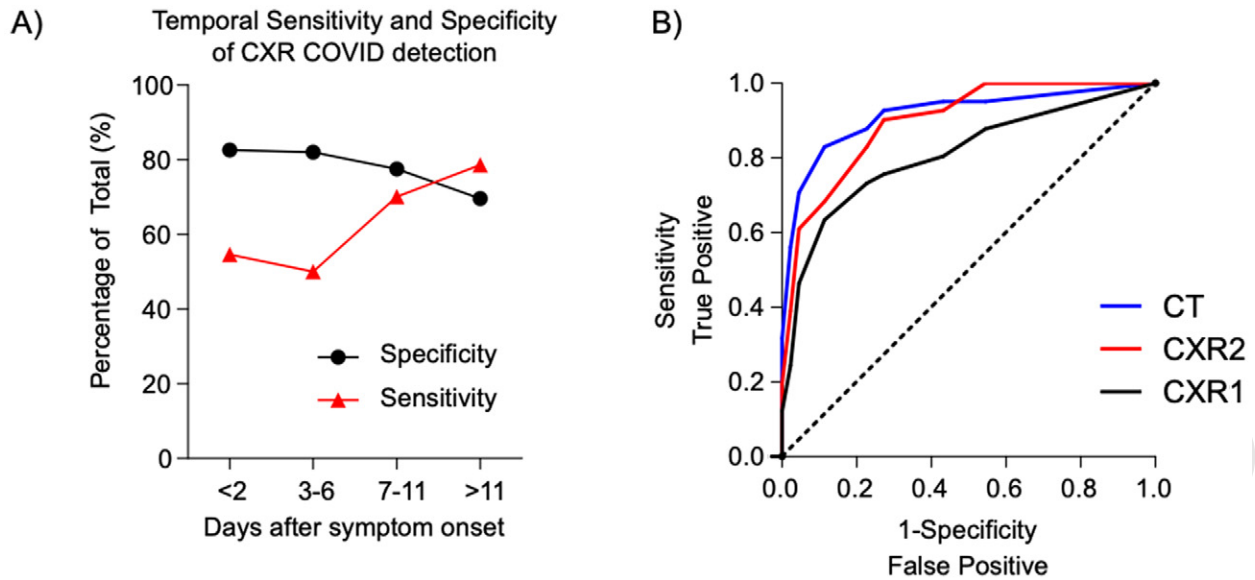
Figure 3. Sensitivity and specificity of chest X-ray and chest CT in COVID-19 infected patients. **(a)** Sensitivity and specificity by days after symptom onset. **(b)** Receiver operating characteristic curves for the 1st chest x-ray (CXR1), 2nd chest x-ray (CXR2), and chest-CT in patients with all three studies (n = 85). Areas under the curve (AUC) of CXR1 = 0.787, AUC of CXR2 = 0.875, and AUC of CT1 = 0.916.

Figure 4. Frequencies of COVID-19+ chest-CT studies with consolidation, ground glass opacities (GGO) alone, or neither, across each time interval.

Tables

Table 1: Patient demographics

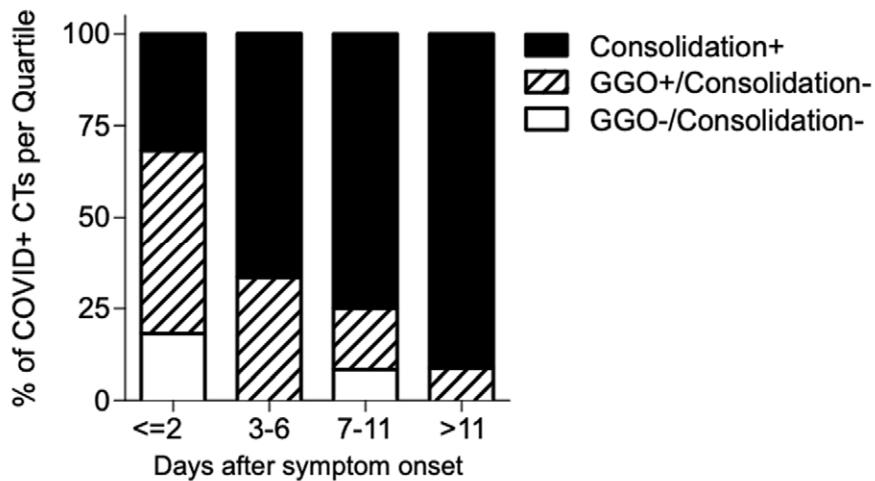| Characteristics	All patients (n = 508)	COVID-19 (+) % (n=254)	COVID-19 (-) % (n=254)
Age, Mean (SD)	56	56 (17.6)	56 (17.5)
Sex - Male	53% (n=268)	53% (n=134)	53% (n=134)
BMI, Median (range)	27.8 (13.3, 71.9)	28.7 (13.3, 71.9)	27.2 (14.5, 66.6)
Ethnicity			
Caucasian-American	47% (n=238)	40% (n=100)	54% (n=138)
African-American	29% (n=148)	28% (n=70)	31% (n=78)
Asian-American	3% (n=16)	2% (n=6)	4% (n=10)
Hispanic-American	10% (n=49)	14% (n=35)	5% (n=14)
Others	11% (n=56)	17% (n=42)	6% (n=14)
Comorbidities			
Hypertension	44% (n=226)	45% (n=114)	44% (n=112)
Diabetes mellitus	26% (n=131)	31% (n=79)	20% (n=52)
Cardiac	28% (n=140)	20% (n=50)	35% (n=90)
Chronic lung disease	24% (n=122)	20% (n=50)	28% (n=72)
Autoimmune disease	6% (n=29)	4% (n=9)	8% (n=20)
Malignancy	15% (n=75)	11% (n=29)	18% (n=46)
SOT/SCT	3% (n=14)	2% (n=6)	3% (n=8)
Chronic Kidney Disease	10% (n=53)	7% (n=18)	14% (n=35)

SOT = Solid Organ Transplant, SCT = Stem Cell Transplant, SD = Standard Deviation

Table 2: Frequency of CXR1 findings in COVID versus non-COVID patients

CXR Finding	Total Frequency Count (%)	Total Frequency Count (%)	Odds Ratio (95% CI)	Adjusted p-value
COVID status	+ (n=249)	- (n=251)		
Peripheral Opacities	52(21%)	10(4%)	6.4 (3.2, 12.8)	<0.001
Interstitial + patchy airspace	127(51%)	66(26%)	2.9 (2.0, 4.2)	<0.001
Normal	52(21%)	106(42%)	0.4 (0.2, 0.5)	<0.001
Normal lung (No Distributive Gradient)	52(21%)	107(43%)	0.4 (0.2, 0.5)	<0.001
Pleural Effusions	13(5%)	43(17%)	0.3 (0.1, 0.5)	<0.001
Atelectasis	20(8%)	62(25%)	0.3 (0.2,0.5)	<0.001
Diffuse Gradient	106(43%)	70(28%)	1.9 (1.3, 2.8)	<0.001
Lobar Consolidation	7(3%)	17(7%)	0.4 (0.2, 1.0)	0.03
Diffuse airspace opacities	15(6%)	6(2%)	2.6 (1.0, 6.9)	0.03
Upper Gradient	7(3%)	2(1%)	3.6 (0.7, 17.5)	0.09
Interstitial Opacities	52(21%)	40(16%)	1.4 (0.9, 2.2)	0.09
Caudal Gradient	84(34%)	72(29%)	1.3 (0.9, 1.9)	0.13

Table 3: Frequency of single CT Findings in COVID versus non-COVID patients

CT Finding	Total Frequency Count (%)	Total Frequency Count (%)	Odds Ratio (95 CI%)	Adjusted p-value
COVID status	+ (n=67)	-(n=102)		
Reverse Halo sign	22 (33%)	5 (5%)	9.5 (3.4, 26.7)	<0.001
GGO	59 (88%)	50 (49%)	7.7 (3.3, 17.7)	<0.001
Craniocaudal Diffuse Gradient	39 (58%)	29 (28%)	3.5 (1.8, 6.7)	<0.001
Normal	2 (3%)	20 (19%)	0.1 (0.03, 0.6)	<0.001
A/P Diffuse Gradient	46 (69%)	48 (47%)	2.5 (1.3, 4.7)	0.004
C/P Central Gradient	0 (0%)	9 (9%)	0 (N/A)	0.009
C/P Peripheral Gradient	25 (37%)	22 (22%)	2.2 (1.1, 4.3)	0.02
Septal Thickening	25 (37%)	22 (22%)	2.2 (1.1, 4.3)	0.02
Craniocaudal Lower Gradient	13 (19%)	34 (33%)	0.5 (0.2, 1.0)	0.03
Consolidation	43 (64%)	51 (50%)	1.8 (1.0, 3.4)	0.04
Pleural Effusions	10 (15%)	26 (25%)	0.5 (0.2, 1.1)	0.07
Nodules	30 (45%)	35 (34%)	1.6 (0.8, 2.9)	0.11
Craniocaudal Upper Gradient	3 (4%)	11 (11%)	0.4 (0.1, 1.4)	0.12
A/P Anterior Gradient	1 (1%)	5 (5%)	0.3 (0.03, 2.6)	0.23
C/P Diffuse Gradient	39 (58%)	53 (52%)	1.3 (0.7, 2.4)	0.26
A/P Posterior Gradient	18 (27%)	32 (31%)	0.8 (0.4, 1.6)	0.32

GGO = Ground glass opacities, A/P = Anterior/Posterior, C/P = Central/Peripheral.

Table 4: Patient demographics and Comorbidities that influence CXR False Negative Rates

Variable	Univariate		Multivariable Regression	
	Odds Ratio	**Adjusted p-value**	**Odds Ratio**	**p-value**
Female Sex	1.6 (1.0, 2.7)	0.04	1.6 (0.9, 2.8)	0.11
Age (per 10 years interval)	0.7 (0.6, 0.9)	<0.001	0.7 (0.6, 0.9)	<0.001
Hypertension	0.6 (0.4, 1.0)	0.04	0.7 (0.4, 1.3)	0.26
Chronic Lung Disease	1.8 (1.0, 3.4)	0.04	1.6 (0.8, 3.2)	0.19
Ethnicity		0.01		0.003
Caucasian-American (reference)	1.0	N/A	1.0	N/A
African-American	1.7 (0.9, 3.1)	0.02	1.5 (0.7, 2.9)	0.02
Asian-American	1.1 (0.2, 5.9)	0.68	1.1 (0.2, 6.5)	0.56
Hispanic-American	0.5 (0.2, 1.0)	0.06	0.3 (0.1, 0.8)	0.02

Supplemental Table 1: Clinical Indications for First Chest X Ray (CXR1)

	Total Frequency Count (%)	**Total Frequency Count (%)**
COVID status	+ (n=249)	- (n=249)
Cough	119 (48%)	102 (41%)
Dyspnea	55 (22%)	49 (20%)
Chest Pain	9 (4%)	22 (9%)
Fever	23 (9%)	20 (8%)
Post-Procedure	17 (7%)	13 (5%)
Hypoxia	6 (2%)	5 (2%)
Others	20 (8%)	34 (14%)

Supplemental Table 2: Clinical Indications for Second Chest X Ray (CXR2)

	Total Frequency Count (%)	Total Frequency Count (%)
COVID status	+ (n=106)	- (n=68)
Dyspnea	16 (15%)	15 (22%)
Cough	22 (21%)	12 (17%)
Line Placement	20 (19%)	12 (17%)
Hypoxia	32 (30%)	6 (9%)
Others	16 (15%)	23 (34 %)

Supplemental Table 3: Clinical Indications for Third Chest X Ray (CXR3)

	Total Frequency Count (%)	Total Frequency Count (%)
COVID status	+ (n=56)	- (n=40)
Dyspnea	8 (14%)	9 (22%)
Line Placement	12 (21%)	8 (20%)
Others	36 (64%)	23 (58%)

Supplemental Table 4: Clinical Indications for Chest CT

	Total Frequency Count (%)	Total Frequency Count (%)
COVID status	+ (n=64)	- (n=99)
Pulmonary Embolism Rule-out	26 (41%)	26 (26%)
Dyspnea	8 (13%)	18 (18%)
Cough	5 (8%)	13 (13%)
Pneumonia	4 (6%)	10 (10%)
Others	21 (33%)	32 (32%)

Supplemental Table 5: Frequency of CXR findings in COVID-19 versus non-COVID-19 patients

CXR Finding	Frequency (%)	Frequency (%)	Odds Ratio (95% CI)	Adjusted p-value
COVID-19 status	+ (n=442)	- (n=364)		
Peripheral Opacities	81 (18%)	10 (3%)	7.9 (4.1, 15.6)	<0.001
Interstitial and airspace opacities	272 (62%)	131 (36%)	2.8 (2.1, 3.8)	<0.001
Diffuse airspace opacities	66 (15%)	22 (6%)	2.7 (1.6, 4.5)	<0.001
Diffuse Gradient	241 (55%)	125 (34%)	2.3 (1.7, 3.0)	<0.001
Lower Gradient	139 (31%)	115 (32%)	1.0 (0.7, 1.3)	0.55
Interstitial Opacities	70 (16%)	61 (17%)	0.9 (0.6, 1.4)	0.40
Lobar Consolidation	23 (5%)	34 (9%)	0.5 (0.3, 0.9)	0.03
Pleural Effusions	45 (10%)	96 (26%)	0.3 (0.2, 0.5)	<0.001
Normal	56 (13%)	120 (33%)	0.3 (0.2, 0.4)	<0.001
Atelectasis	36 (8%)	98 (27%)	0.2 (0.2, 0.4)	<0.001
Upper Gradient	7 (2%)	1 (1%)	1.9 (0.5, 7.5)	0.26
Normal Distributive gradient	55 (12%)	121 (33%)	0.3 (0.2, 0.4)	<0.001
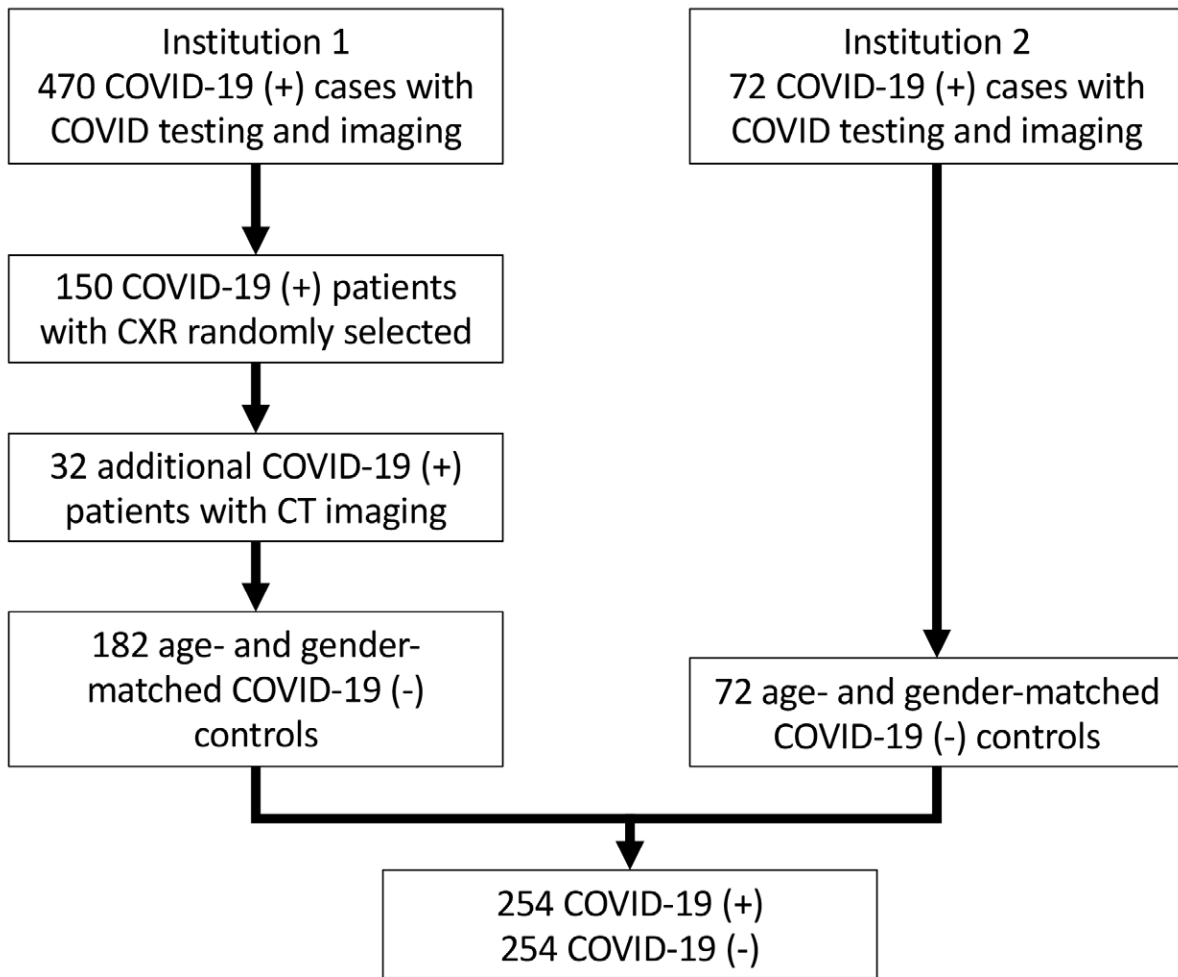
Supplementary Figure 1: CONSORT Diagram for inclusion of COVID-19 positive cases and COVID-19 negative controls across two different institutions.