

RESEARCH

Open Access



Assessment of interobserver reliability and predictive values of CT semiquantitative and severity scores in COVID lung disease

Dhilip Andrew, Karthik Shyam* , Soumya Cicilet and Jovis Johny

Abstract

Background: The coronavirus disease (COVID-19), caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), and first reported in December 2019 at Wuhan, China, has since then progressed into an ongoing global pandemic. The primary organ targeted by the virus is the pulmonary system, leading to interstitial pneumonia and subsequent oxygen dependency and morbidity. Computed tomography (CT) has been used by various centers as an imaging modality for the assessment of severity of lung involvement in individuals. Two popular systems of scoring lung involvement on CT are CT semiquantitative score (SQ) and CT severity score (CT-SS), both of which assess extent of pulmonary involvement by interstitial pneumonia and are partly based upon subjective evaluation. Our cross-sectional observational study aims to assess the interobserver reliability of these scores, as well as to assess the statistical correlation between the respective CT scores to severity of clinical outcome.

Results: Both the SQ and CT-SS scores showed an excellent interobserver reliability (ICC 0.91 and 0.93, respectively, $p < 0.05$). The CT-SS was marginally more sensitive (99.2%) in detecting severe COVID pneumonia than SQ (86.5%). The positive predictive value of SQ (98.3%) is more than CT-SS (78%) for detecting severe disease. The similarity of interobserver reliability obtained for both scores reiterates the respective cutoff CT scores proposed by the above systems, as 18 for SQ and 19.5 for CT-SS.

Conclusion: Both the SQ and CT-SS scores display excellent interobserver reliability. The CT-SS was more sensitive in detecting severe COVID pneumonia and may thus be preferred over the SQ as an initial radiological tool in predicting severity of infection.

Keywords: COVID, Pneumonia, HRCT, Score, Reliability

Background

The coronavirus disease (COVID-19) is caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). The outbreak was first reported in December 2019 at Wuhan, China, since then the disease has progressed into an ongoing global pandemic [1]. As of December 15th 2020, there have been 71,581,532 confirmed cases and 1,618,374 deaths [2]. The disease affects patients of all ages; however, they can progress rapidly in elderly patients and

patients with comorbidities (such as hypertension, chronic heart disease, chronic obstructive pulmonary disease, and diabetes). Elderly patients and those with comorbidities can present with multiple systemic complications secondary to the disease [3, 4]. The current gold standard for diagnosis of COVID-19 is real-time polymerase chain reaction (RT-PCR) [5]. Role of radiological imaging in screening and diagnosis of COVID is much debated, with recommendations against and supportive of chest computed tomogram (CT) as first line of screening. However, according to the guidelines of European Society of Radiology and European Society of Thoracic Imaging the

* Correspondence: dr.karthikshyam@gmail.com

Department of Radiology, St. John's Medical College Hospital, Sarjapur Road, Bangalore, Karnataka 560034, India



© The Author(s). 2021 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

limited role of chest CT in resource limited set up is recognized [6, 7]. Chest imaging plays an important role in working up and staging of COVID-19: it can help stratify cases of COVID-19 based on imaging findings and find patients most at risk of adverse clinical outcomes [7–9]. Chest CT has a higher sensitivity (86–98%) and lower false-negative rate compared to RT-PCR [10–12].

The more commonly reported CT chest findings in COVID-19 patients are ground glass opacities (GGO), reticular opacities, and consolidation. Less common CT findings were subpleural lines, crazy paving sign (ground glass opacities superimposed with interlobular lines and septal thickening), and bronchial wall thickening. Atypical CT findings are pleural effusion and mediastinal lymphadenopathy [6, 9, 13–16]. The predominant distribution of findings are bilateral, peripheral, subpleural, posterior, and basal in distribution [9, 14]. The CT findings observed in COVID-19 are not specific and overlap with a number of other conditions [12, 17, 18]. Given the need for a common consensus and language for reporting the Chest CT findings in COVID, CT reporting guidelines by RSNA (Radiology Society of North America) and CO-RADS by Dutch Radiological Society were proposed. These guidelines provided the radiologist with terminology and language required for reporting the findings and thus helped improving interobserver agreement. The interobserver agreement for RSNA score and CO-RADS were excellent (Fleiss kappa value of 0.871 and 0.876) [18].

According to study by Jiong et al., chest CT could be used to evaluate the clinical severity of the disease since there were correlation between lung involvement, laboratory markers, and clinical symptoms [9]. Various CT scoring systems have been proposed to assess the lung involvement, the two most commonly used scores being the semiquantitative (SQ) score and CT severity score (CT-SS) [19, 20]. CT scoring systems have been used to predict the clinical outcome like mortality and chart the prognosis of the patient [21, 22]. Given the vital role of CT scoring system in monitoring the clinical prognosis of the COVID-19 patient, there is need for a CT score with good interobserver agreement and correlation with clinical severity.

Aims and objectives

To compare CT-SS and SQ scores and find the score with the least interobserver variation for determining the lung involvement in COVID-19 pneumonia, and to determine the correlation between lung involvement and clinical severity of COVID pneumonia.

Method

A cross-sectional study was undertaken at a tertiary care referral academic center designated for management of

COVID-19 patients. The study was conducted between June 2020 and December 2020. Institutional ethical clearance was obtained and informed consent was waived.

Patient selection

A total of 120 adults (≥ 18 years old) hospitalized with laboratory confirmed COVID-19 infection were included in the study. Inclusion criteria included patients who have newly tested positive for COVID-19 (RT-PCR) at the time of admission or after admission, and had a CT scan of the chest performed during admission. Exclusion criteria included patients preliminarily treated for COVID-19 in another hospital and individuals with reinfection with COVID-19.

Clinical data collection

Clinical and laboratory data such as demographic details, symptoms at presentation, examination findings, arterial blood gas findings, and CT chest findings were collected. According to the following criteria, patients were clinically categorized into mild and severe pneumonia according to guidelines laid by Chinese Center for Disease Control [23]. Patients were categorized with *mild disease* (mild symptoms without dyspnea, respiratory frequency $< 30/\text{min}$; blood oxygen saturation (SpO_2) $> 93\%$; $\text{PaO}_2/\text{FiO}_2$ ratio ≥ 300 mmHg), *severe disease* (dyspnea, respiratory frequency $\geq 30/\text{min}$, $\text{SpO}_2 \leq 93\%$, $\text{PaO}_2/\text{FiO}_2$ ratio < 300 mmHg, and/or lung infiltrates on X-ray $> 50\%$ within 24 to 48 h), and *critical disease* (patients with adult respiratory distress syndrome or respiratory failure/acute respiratory distress syndrome (ARDS), septic shock, and multiple organ dysfunction (MOD) or multi-organ failure (MOF). The patients in the critical disease group were categorized under the severe disease group for the purpose of our study.

Image acquisition

Image acquisition was carried out using a helical 64 multi-slice CT scanner. Images were acquired in a single breath hold spanning entire chest from the diaphragmatic domes to lung apices. CT scan parameters were as follows: X-ray tube parameters 120 kVp, 250 mAs; rotation time 0.5 s; pitch 1.0; section thickness 5 mm; intersection space 5 mm. Images were reconstructed to 0.625 mm thickness sections using the soft tissue and lung kernels. Images were analyzed using standard viewing software.

Image analysis

Images were separately evaluated by two radiologists with 4 and 10 years of experience in thoracic CT interpretation, while being blinded to the clinical details. Findings such as ground glass opacities, consolidation,

and crazy paving were recorded using the standard nomenclature defined by the Fleischner Society glossary [24]. CT-SS and SQ scores for each patient were calculated. In case of CT-SS, lung opacities in all of the 20 segments of lung were evaluated and a score of either 0, 1, or 2 were given if 0%, less than 50%, or 50% or more of each segment was involved, respectively. Lung segments were determined according to the system utilized by Yang et al. [20], where each lobe is divided into its constituent segments, and the lingular segment being divided into superior and inferior subsegments. The final CT-SS score is the sum of these individual scores with a possible range from 0 to 40 [20] (Fig. 1). A score of 19 or more was categorized as radiologically severe. In case of SQ score, lung opacities in all the 5 lobes of the lung were evaluated and scored individually: 0 if no involvement, 1 if < 5% involved, 2 if 5–25% involved, 3 if 26–50% involved, 4 if 51–75% involved and 5 if > 75% involved. The SQ score is the sum of individual score of each lobe, with a possible range of 0 to 25 (Fig. 2). SQ score of 18 and more was categorized as radiologically severe.

Statistical analysis

Statistical analysis was performed using the software SPSS (version 24, IBM). Quantitative data are presented as mean, standard deviation, and ranges. Qualitative data are presented as percentages of total. Kappa statistics was determined for each CT scoring system based on categorical variable outcomes (mild vs. severe). Interclass correlation coefficient was used for each scoring system utilizing the continuous variables (absolute value of CT score). *P* value of less than 0.05 was chosen to indicate statistical significance.

Results

Out of the 115 patients admitted with confirmed COVID pneumonia between June 2020 and December 2020, 57% patients ($n = 66$) were classified as severe. Twenty-six percent ($n = 30$) underwent clinical deterioration resulting in mortality. The median age of the study group was 46 years (IQR 19–73 years) with 63% ($n = 72$) being males. As the initial presenting complaint, 36% ($n = 41$) with cough, 42% ($n = 48$) with fever, and 22% ($n = 25$) with breathlessness.

Sensitivity for the SQ score varied from 85 to 88% (mean 86.5%) while the specificity ranged from 96 to 100%. Sensitivity for the CT-SS test was noted to range from 98.5 to 100% (mean 99.2%) while the specificity ranged from 59 to 65%.

The positive predictive value for SQ test ranged from 96.7 to 100% (mean 98.3%), while the negative predictive value ranged from 83 to 85.5% (mean 85.7%).

The positive predictive value ranged from 77 to 79% (mean 78%), while the negative predictive value ranged from 97 to 100% (mean 98.5%) for the CT SS test.

While adopting a random effects model, there was excellent reliability between the two observers for both SQ and CT-SS, with an agreement of 91.3% and 88%, respectively.

Kappa statistics values for interobserver correlation of 91% and 93% were seen for SQ and CT-SS scores respectively after categorization of the patients into mild and severe COVID cases, again indicating an excellent inter-observer agreement. The extent of this agreement is similar to that found by Jiang et al. [25] for SQ scores (0.91), and by Yang et al. for CT-SS (0.936), thus reiterating the cutoff scores for diagnosing severe COVID pneumonia as 18 for SQ, and 19.5 for CT-SS.



Fig. 1 a–c 45-year-old female with clinically mild COVID and SQS of 8, d–f 52-year-old male with clinically severe COVID and SQS of 22

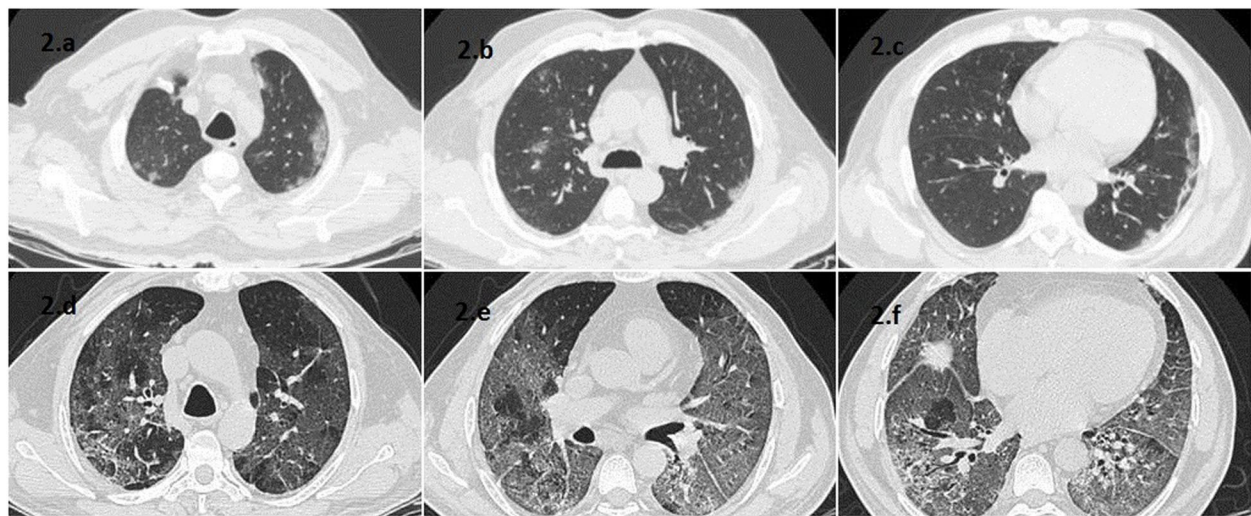


Fig 2 a-c 50-year-old male with clinically mild COVID pneumonia and CTSS of 11, **d-f** 45-year-old male with clinically severe COVID pneumonia and CTSS of 32

Discussion

The rapidly evolving global scenario with respect to the current pandemic caused by SARS-CoV-2 has necessitated the orchestrated efforts of various specialties in the diagnosis and treatment of the disease, and radiology holds an essential position in this regard. Though the initial diagnosis is made by way of tests such as the rapid antigen test (RAT) or real-time polymerase chain reaction (rt-PCR), chest radiographs are used for initial severe presentations of pneumonia [10]. CT scan of the chest is reserved primarily for complicated or unresolved pneumonia, or other suspected thoracic pathologies, and is more sensitive to recognizing the pulmonary manifestations of COVID lung disease in comparison to radiographs [26, 27]. In view of this, the proposed CT chest scoring systems, the semiquantitative (SQ) score proposed by Francone et al. [21] and the CT severity score, proposed by Yang et al. [20] have been in worldwide use to help diagnose and prognosticate COVID pneumonia.

In view of the relative nascency of imaging standardisation of COVID lung disease, and the subsequent differing systems of scoring used, it would be prudent to assess the extent of interobserver variability of the above two systems, and to help determine which, if any, more accurately represents the clinical severity and prognosis. To our knowledge, such a study comparing different CT scoring systems for COVID lung disease has not yet been performed.

In our sample population of 115 patients, a majority were males (63%), and percentages of individuals presenting with cough (36%) and fever (42%). Among two experienced radiologists who performed the blinded scoring, sensitivity for diagnosis of severe COVID

pneumonia with respect to the SQ score was 85–88%, and 98.5–100% for the CT-SS. This implies a marginally higher sensitivity for CT-SS in identifying severe COVID lung disease. Specificity for recognizing severe COVID lung disease was similar for both scores. The interobserver reliability for both scores was excellent: the interclass correlation when comparing continuous variables (absolute value of score) were 91.3% and 88% for SQ and CT-SS respectively. The interobserver reliability determined via Kappa statistics for categorical variables (mild-moderate vs. severe-critical) were 91% and 93% for SQ and CT-SS, respectively. Limitations of this study include the limited sample size, and subjectivity in assigning scores. This is an issue that has been raised in previous studies as well [28].

Conclusion

Our study shows that there exists excellent interobserver agreement for both the SQ and CT-SS scores with regard to the diagnosis of severe COVID lung disease. CT-SS displays marginally higher sensitivity in identifying severe COVID, and may thus be preferred by radiologists as a standard screening tool while assessing CT. The rapid and uniform adoption of this system of scoring may assist in the effective communication of disease severity to referring clinicians as well as between radiologists themselves, for both first-time assessments and for comparison purposes with previous studies to evaluate disease progression. Also, the development of scoring systems based on CT imaging of the chest and the wealth of promising data obtained with regard to correlation of such scores with clinical outcome and laboratory parameters open up new possibilities for

further development of novel scoring systems for other infective/inflammatory lung diseases in the future.

Abbreviations

COVID-19: Coronavirus disease 2019; SARS-CoV-2: Severe acute respiratory syndrome coronavirus 2; RT-PCR: Real-time polymerase chain reaction; CT: Computed tomogram; SQ: Semiquantitative score; CT-SS: CT severity score; ARDS: Acute respiratory distress syndrome; MOD: Multiple organ dysfunction; MOF: Multi-organ failure

Acknowledgements

The authors would like to thank all those involved in the design, execution, and interpretation of this study.

Authors' contributions

DA and KS contributed towards the design of the study, and selection of appropriate cases. DA and SC contributed to scoring the CT images. JJ and KS contributed towards the statistical analysis. KS and DA contributed towards the development of the manuscript. All authors have read and approved of the final manuscript.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. There were no sources of funding for this work other than departmental resources.

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate

Requisite verbal ethical clearance was provided by the Institutional Ethical Clearance Board (IERB) of St. John's Medical College (number not available). Need for patient consent was waived by the above review board in view of this being a retrospective study.

Consent for publication

Not applicable.

Competing interests

The authors declare that they do not have any competing interests.

Received: 25 January 2021 Accepted: 2 June 2021

Published online: 18 June 2021

References

- Ng JY (2020) Global research trends at the intersection of coronavirus disease 2019 (COVID-19) and traditional, integrative, and complementary and alternative medicine: a bibliometric analysis. *BMC Complement Med Ther* 20(1). <https://doi.org/10.1186/s12906-020-03151-8>
- World Health Organization. Coronavirus disease (COVID-19). Published online December 15, 2020. <https://www.who.int/emergencies/diseases/novel-coronavirus-2019>
- Nokhodian Z, Ranjbar M, Nasri P, Kassaian N, Shoaee P, Vakili B, Rostami S, Ahangarzadeh S, Alibakhshi A, Yarian F, Javanmard SH, Ataei B (2020) Current status of COVID-19 pandemic; characteristics, diagnosis, prevention, and treatment. *J Res Med Sci* 25(1):101. https://doi.org/10.4103/jrms.JRMS_476_20
- Singhal T (2020) A review of coronavirus disease-2019 (COVID-19). *Indian J Pediatr* 87(4):281–286. <https://doi.org/10.1007/s12098-020-03263-6>
- Alpdagtas S, Ilhan E, Uysal E, Sengor M, Ustundag CB, Gunduz O (2020) Evaluation of current diagnostic methods for COVID-19. *APL Bioeng* 4(4): 041506. <https://doi.org/10.1063/5.0021554>
- Fields BKK, Demirjian NL, Dadgar H, Gholamrezanezhad A Imaging of COVID-19: CT, MRI, and PET. *Semin Nucl Med*. Published online November 2020. <https://doi.org/10.1053/j.semnuclmed.2020.11.003>
- Revel M-P, Parkar AP, Prosch H et al (2020) COVID-19 patients and the radiology department – advice from the European Society of Radiology (ESR) and the European Society of Thoracic Imaging (ESTI). *Eur Radiol* 30(9): 4903–4909. <https://doi.org/10.1007/s00330-020-06865-y>
- Kooraki S, Hosseiny M, Myers L, Gholamrezanezhad A (2020) Coronavirus (COVID-19) outbreak: what the Department of Radiology Should Know. *J Am Coll Radiol* 17(4):447–451. <https://doi.org/10.1016/j.jacr.2020.02.008>
- Wu J, Wu X, Zeng W, Guo D, Fang Z, Chen L, Huang H, Li C (2020) Chest CT findings in patients with coronavirus disease 2019 and its relationship with clinical features. *Invest Radiol* 55(5):257–261. <https://doi.org/10.1097/RLI.0000000000000670>
- Ai T, Yang Z, Hou H, Zhan C, Chen C, Lv W, Tao Q, Sun Z, Xia L. Correlation of chest CT and RT-PCR testing for coronavirus disease 2019 (COVID-19) in China: a report of 1014 cases. *Radiology*. 2020;296(2):E32–40.
- Guan W, Ni Z, Hu Y, Liang WH, Ou CQ, He JX, Liu L, Shan H, Lei CL, Hui DSC, du B, Li LJ, Zeng G, Yuen KY, Chen RC, Tang CL, Wang T, Chen PY, Xiang J, Li SY, Wang JL, Liang ZJ, Peng YX, Wei L, Liu Y, Hu YH, Peng P, Wang JM, Liu JY, Chen Z, Li G, Zheng ZJ, Qiu SQ, Luo J, Ye CJ, Zhu SY, Zhong NS, China Medical Treatment Expert Group for Covid-19 (2020) Clinical characteristics of coronavirus disease 2019 in China. *N Engl J Med* 382(18): 1708–1720. <https://doi.org/10.1056/NEJMoa2002032>
- Fang Y, Zhang H, Xie J, Lin M, Ying L, Pang P, Ji W (2020) Sensitivity of chest CT for COVID-19: comparison to RT-PCR. *Radiology*. 296(2):E115–E117. <https://doi.org/10.1148/radiol.2020200432>
- Zheng Y, Wang L, Ben S. Meta-analysis of chest CT features of patients with COVID-19 pneumonia. *J Med Virol*. Published online July 11, 2020. doi: <https://doi.org/10.1002/jmv.26218>
- Song F, Shi N, Shan F, Zhang Z, Shen J, Lu H, Ling Y, Jiang Y, Shi Y (2020) Emerging 2019 novel coronavirus (2019-nCoV) pneumonia. *Radiology*. 295(1):210–217. <https://doi.org/10.1148/radiol.2020200274>
- Carotti M, Salaffi F, Sarzi-Puttini P, Agostini A, Borgheresi A, Minorati D, Galli M, Marotto D, Giovagnoni A (2020) Chest CT features of coronavirus disease 2019 (COVID-19) pneumonia: key points for radiologists. *Radiol Med (Torino)* 125(7):636–646. <https://doi.org/10.1007/s11547-020-01237-4>
- Zhao W, Zhong Z, Xie X, Yu Q, Liu J (2020) Relation between chest CT findings and clinical conditions of coronavirus disease (COVID-19) pneumonia: a multicenter study. *Am J Roentgenol* 214(5):1072–1077. <https://doi.org/10.2214/AJR.20.22976>
- Huang P, Liu T, Huang L, Liu H, Lei M, Xu W, Hu X, Chen J, Liu B (2020) Use of chest CT in combination with negative RT-PCR assay for the 2019 novel coronavirus but high clinical suspicion. *Radiology*. 295(1):22–23. <https://doi.org/10.1148/radiol.2020200330>
- O' Neill SB, Byrne D, Müller NL, et al. Radiological Society of North America (RSNA) expert consensus statement related to chest CT findings in COVID-19 versus CO-RADS: comparison of reporting system performance among chest radiologists and end-user preference. *Can Assoc Radiol J*. Published online November 3, 2020:084653712096891. doi:10.1177/0846537120968919
- Pan F, Ye T, Sun P, Gui S, Liang B, Li L, Zheng D, Wang J, Hesketh RL, Yang L, Zheng C (2020) Time course of lung changes at chest CT during recovery from coronavirus disease 2019 (COVID-19). *Radiology*. 295(3):715–721. <https://doi.org/10.1148/radiol.2020200370>
- Yang R, Li X, Liu H, Zhen Y, Zhang X, Xiong Q, Luo Y, Gao C, Zeng W (2020) Chest CT severity score: an imaging tool for assessing severe COVID-19. *Radiol Cardiothorac Imaging* 2(2):e200047. <https://doi.org/10.1148/ryct.2020.00047>
- Francone M, Iafrate F, Masci GM, Coco S, Cilia F, Manganaro L, Panebianco V, Andreoli C, Colaiacono MC, Zingaropoli MA, Ciardi MR, Mastroianni CM, Pugliese F, Alessandri F, Turriziani O, Ricci P, Catalano C (2020) Chest CT score in COVID-19 patients: correlation with disease severity and short-term prognosis. *Eur Radiol* 30(12):6808–6817. <https://doi.org/10.1007/s00330-020-07033-y>
- Abbasi B, Akhavan R, Ghamari Khameneh A, et al. Evaluation of the relationship between inpatient COVID-19 mortality and chest CT severity score. *Am J Emerg Med*. Published online September 2020. doi:<https://doi.org/10.1016/j.ajem.2020.09.056>
- Wu Z, McGoogan JM (2020) 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*. 323(13):1239–1242. <https://doi.org/10.1001/jama.2020.2648>
- Hansell DM, Bankier AA, MacMahon H, McLoud TC, Müller NL, Remy J (2008) Fleischner Society: Glossary of Terms for Thoracic Imaging. *Radiology*. 246(3):697–722. <https://doi.org/10.1148/radiol.2462070712>

25. Jiang Y, Guo D, Li C, Chen T, Li R (2020) High-resolution CT features of the COVID-19 infection in Nanchong City: Initial and follow-up changes among different clinical types. *Radiol Infect Dis* 7(2):71–77. <https://doi.org/10.1016/j.jrid.2020.05.001>
26. Erturk SM (2020) CT Is not a screening tool for coronavirus disease (COVID-19) pneumonia. *Am J Roentgenol* 215(1):W11
27. Lin Y, Zhao W, Liu J (2020) Reply to “CT Is not a screening tool for coronavirus disease (COVID-19) pneumonia”. *Am J Roentgenol* 215(1):W12
28. Saeed GA, Gaba W, Shah A, Al Helali AA, Raidullah E, Al Ali AB, Elghazali M, Ahmed DY, Al Kaabi SG, Almazrouei S. Correlation between Chest CT Severity Scores and the Clinical Parameters of Adult Patients with COVID-19 pneumonia. *Radiol Res Pract*. 2020. p. 2021.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

Submit your next manuscript at ▶ [springeropen.com](https://www.springeropen.com)
