Published online 2021 July 20 **Original Article**

Evaluation of Chest CT Findings using the Reporting and Data System of Patients with Suspected COVID-19 Infection

Ibrahim Halil Sever[1,*], Bahattin Ozkul[1], Bedriye Koyuncu Sokmen[1] and Nagihan Inan Gurcan[1]

[1] Department of Radiology, Florence Nightingale Hospital, Demiroğlu Bilim University, Istanbul, Turkey

*** *Corresponding author:*** Ibrahim Halil Sever, Department of Radiology, Florence Nightingale Hospital, Demiroğlu Bilim University, Istanbul, Turkey. Tel: +905532040366; Email: halilsever4022@gmail.com

Received 2021 February 28; **Revised** 2021 May 25; **Accepted** 2021 August 03.

Abstract

Background: The simultaneous interpretation of computed tomography (CT) scans performed on patients with suspected clinical signs of coronavirus disease 2019 (COVID-19) or a history of contact may accelerate patient isolation, particularly during the peak of the pandemic. The use of an appropriate scoring system can lead to the conveyance of the findings in a more understandable way and the elimination of differences in interpretations.

Objectives: This study aimed to evaluate the diagnostic performance of the coronavirus disease 2019 (COVID-19) imaging reporting and data system (CO-RADS) in admitted patients with suspected COVID-19 infection.

Methods: This retrospective study included all patients admitted to our hospital with COVID-19 pneumonia suspicion within March 20-May 15, 2020, who were examined by both CT and real-time reverse transcription polymerase chain reaction (rRT-PCR) at initial presentation. Four radiologists, who were blinded to the rRT-PCR results and medical history, assessed all images independently and classified the CT findings according to the CO-RADS previously defined. Diagnostic value of the scoring system and interobserver agreement in rRT-PCR positive-negative groups and for CO-RADS 1-5 were evaluated.

Results: In this study, 274 (153 men and 121 women; 48.8±17.3 years) rRT-PCR positive and 437 (208 men and 229 women; 49.0±19.5 years) rRT-PCR negative individuals were included. It was found that CO-RADS had a good diagnostic performance with area under the receiver operating characteristic roc curve of 0.857. The sensitivity, specificity, positive predictive value, negative predictive value, and accuracy were obtained at 81.9%, 89.4%, 75.7%, 92.5%, and 84.8%, respectively. The interobserver agreement of four radiologists in CO-RADS 1 and 5 was substantial to almost perfect according to the kappa values. Other CO-RADS scores showed a fair to moderate agreement. The interrater agreement was slightly higher in the PCR (-) patient group than in the positive one.

Conclusion: In conclusion, CO-RADS was a successful scoring system for distinguishing highly suspicious cases in terms of COVID-19 infection lung involvement, showing high interobserver agreement.

Keywords: CO-RADS, COVID-19, CT, SARS-CoV-2 PCR, Scoring

1. Background

In December 2019, an outbreak with respiratory symptoms caused by a novel coronavirus (i.e., severe acute respiratory syndrome coronavirus 2 [SARS-CoV-2]) appeared in Wuhan, China. Due to its high rate of human-to-human transmission, it spread almost all countries in a short while, which made the World Health Organization declare it as a pandemic on March 11, 2020 (1). Now at the end of May 2021, almost 170 million people have been infected and about 3.5 million people have passed away because of this pandemic (1).

Although the most common clinical symptoms are fever, cough, and dyspnea, non-specific symptoms, such as headache, muscle soreness, fatigue, nausea, diarrhea, loss of smell and taste may appear within 2-14 days after exposure (2). The gold standard diagnostic method for coronavirus disease 2019 (COVID-19) is real-time reverse transcription-polymerase chain reaction (rRT-PCR). However, rRT-PCR is not useful in rapid assessment with a long turnaround time requiring hours, and unclear sensitivity values vary between 42% and 83% depending on sample quality, viral load, and duration of symptoms (3-6). Accordingly, RT-PCR testing should be repeated in patients with a negative initial result, causing a risk for infecting a larger population.

Computed tomography (CT) can be performed rapidly and has a high accuracy rate in clinically suspected cases with uncertain laboratory test results and asymptomatic individuals with known exposure (7, 8). Although some typical findings of COVID-19 infection can be easily detected on CT examination, imaging features overlapping with some conditions, such as different viral diseases and interstitial lung diseases, may also be encountered, causing erroneous interpretations. In addition, since the pandemic is highly recent, CT findings described in the literature are open to different interpretations among radiologists according to their level of experience. For the current reasons, CT findings cannot give clear information in some cases in the diagnosis of COVID-19. These situations make it difficult to

communicate between radiologists and referring physicians. Clinicians and radiologists need a more clear statement about what the observed findings mean to provide isolation as quickly as possible. Therefore, the introduction of standardized reporting systems for patients with suspected COVID-19 is needed to improve communication between physicians and obtain more quantitative data instead of qualitative expressions.

To standardize the interpretation of the chest CT images, such scoring systems have been developed by different groups. The British Society of Thoracic Imaging (BSTI) proposed guidance for radiologists and classified chest CT findings as Classic, Probable, Indeterminate, and Non-COVID (9). The Radiological Society of North America (RSNA) Expert Consensus Statement suggests a classification for COVID-19 pneumonia made up of four categories, namely negative for pneumonia, atypical appearance, indeterminate appearance, and typical appearance (10). Another group of researchers called their systems Coronavirus Disease Reporting and Data System (COVID-RADS) and classified CT findings as COVID-RADS 0, 1, 2A, 2B, and 3 (11). The Dutch Radiological Society (NVvR) described CO-RADS (COVID-19 Reporting and Data System) that correlated with the scoring systems previously defined and classified chest CT findings according to scores ranging from 0 to 6 (2).

2. Objectives

This study aimed to evaluate the diagnostic performance of CO-RADS not only in rRT-PCR positive patients but also in all patients who apply with the suspicion of COVID-19 infection and undergo chest CT.

3. Methods

This retrospective study was approved by the Institutional Review Board (Approved date/number: 13.10.2020 / 2020-19-04). Written informed consent was waived by the committee.

3.1.Patients

The statistical population of this study consisted of 711 patients (274 rRT-PCR positive and 437 rRT-PCR negative) undergoing non-contrast chest tomography due to clinical suspicion of COVID-19 infection (i.e., fever of >37.5°C, cough, and/or shortness of breath) within March 20-May 15, 2020. Nasopharyngeal and oropharyngeal swab specimens were taken from all patients. Patients who showed at least one positive RT-PCR were considered to be positive for COVID-19; otherwise, if the CT findings were suggestive of COVID-19, repeated RT-PCR testing was performed up to a maximum of three times. Patients with initial negative RT-PCR and negative CT findings underwent a 14-day follow-up. All CT scans were performed within the first 3 days after blood sample collection. Patients whose CT images could not be evaluated clearly due to artefacts were excluded from the study (CO-RADS 0). Furthermore, CO-RADS 6 was not used in this study since the evaluation of the images was conducted completely without the laboratory result.

3.2. Imaging Technique

All CT examinations were performed by a 16-slice multidetector-row CT scanner (Somatom Go Now, Siemens Healthcare, Erlangen, Germany). The scanning parameters were as follows: 120 kV; 250 mAs; rotation time, 0.8-second pitch, 1.5. All data were reconstructed at 1.0 mm slice thickness, 512×512 matrix size and a sharp reconstruction kernel (KernelBr64). Acquisitions were performed in the supine position, during a deep inspiration breath-hold and without contrast media administration.

3.3. Image Analysis

Four radiologists (with 10, 12, 12, and 18 years of experience in chest imaging) who were blinded to the rRT-PCR results and medical history, assessed all images independently using a dedicated workstation (General Electric Company, Centricity RIS 6, version 6.0.7.7). All raters had at least 100 CT interpretation experience with COVID-19 lung involvement and each rater practiced separately on the images containing 10 CT samples from each patient group (totally 50 CT samples different from the study group) from CO-RADS 1 to 5. Radiologists evaluated and classified the chest CT findings based on the scoring system created by the NVvR for COVID-19 infection (2). According to this system, pulmonary findings of COVID-19 infection were classified into seven different groups. Accordingly, CO-RADS 0 was defined as if the scan was incomplete or insufficient quality due to artefacts (Figure 1a), CO-RADS 1 was described as a very low level of suspicion (Figure 1b), CO-RADS 2 as a low level of suspicion (Figure 1c), CO-RADS 3 as equivocal findings (Figure 1d), CO-RADS 4 as a high level of suspicion (Figure 1e), and CO-RADS 5 as a very high level of suspicion (Figure 1f) for pulmonary involvement of COVID-19 infection. Finally, CO-RADS 6 was defined for the chest CT findings of the patient whose COVID-19 infection was proven by the rRT-PCR test.

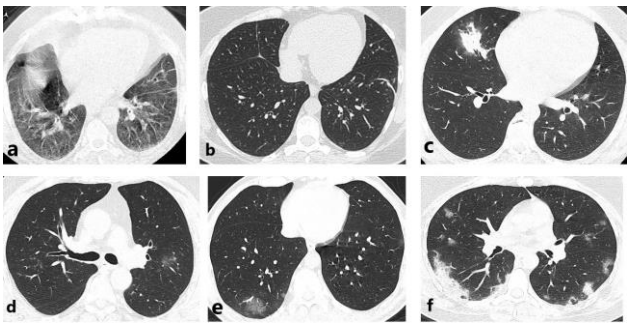

Figure 1. Figure subunits representing illustration of the CO-RADS
Axial (a-f) non-contrast chest CT images: CT scan (a) that cannot be evaluated due to motion artifacts (CO-RADS 0), (b) pleuroparenchymal band located in the left lung (CO-RADS 1, very low level of suspicion), (c) peribronchial consolidation located in the right lung (CO-RADS 2, low level of suspicion), (d) ground-glass opacity located along the peribronchial zone (CO-RADS 3, ındeterminate), (e) ground-glass opacity located in the peripheric zone of the left lung (CO-RADS 4, high level of suspicion), and (f) bilateral ground-glass opacities with crazy paving pattern located mostly in the peripheral zone (CO-RADS 5, very high level of suspicion)

3.4. Statistical Analysis

All statistical analyses were performed in SPSS.25 software for macOS. The fitness of numeric data set to normal distribution was determined by the Kolmogorov-Smirnov test, and the measurement data were expressed as mean±standard deviation (SD). The CT findings of patients between PCR positive and negative groups were compared with the χ^2 test. The Mann-Whitney U test was performed to assess differences in age between the two groups. Fleiss kappa (κ) with %95 confidence intervals (CIs) was used to evaluate the interobserver agreement and to determine the level of agreement between all observers. Based on the description provided by Landis and Koch (12), κ values were divided into groups, including slight agreement (0.01-0.20), fair agreement (0.21-0.40), moderate agreement (0.41-0.60), substantial agreement (0.61-0.80), and almost perfect agreement (0.81-1). A receiver operating characteristics (ROC) curve was calculated, and the area under the ROC curve (AUC) was used to assess the diagnostic performance of CO-RADS relative to a positive rRT-PCR test. Moreover, for each radiologist, the highest Youden index was calculated to select the optimal threshold to discriminate between COVID-19 positive and negative patients, and the corresponding sensitivity, specificity, positive predictive value, and negative predictive value were calculated. The p-value of < 0.05 was considered statistically significant.

4. Results

In our retrospective cohort of 753 patients, 711 individuals were included in the study. Among the participants, 274 patients (153 men and 121 women; 48.8±17.3 years) were rRT-PCR positive and 437 patients (208 men and 229 women; 49.0±19.5 years) were rRT-PCR negative. In this study, 28 patients were excluded when the time interval between chest CT and the rRT-PCR assay was longer than 3 days. Furthermore, 14 patients were excluded because of the inadequate diagnostic quality of imaging due to artefacts.

All individuals had nasopharyngeal and oropharyngeal swab samples. Out of 274 rRT-PCR positive patients, 82 and 24 patients had 2 and 3 tests, respectively. In the same vein, 176 of 437 rRT-PCR negative patients had 2 tests and the others had 1 test.

The PCR positive and PCR negative groups were compared in terms of demographic characteristics and imaging findings. In correlation with the data accumulation defined so far, a significant difference was observed between the two groups in terms of ground glass-opacity (GGO), vascular enlargement, crazy paving pattern, reverse halo sign, and hospitalization. However, as can be predicted,there was no significant difference in other variables (Table 1).

Table 1. Demographics and chest computed tomography findings of patients

Characteristics	PCR +	PCR -	p (2-tailed)
Age (years)*	48.8±17.3	49.0±19.546 (32-64)	0.776*
Gender			0.033***
Male	153 (55.8)	208 (47.6)	
Female	121 (44.2)	229 (52.4)	
Hospitalization	70 (25.5)	70 (16)	0.002***
GGO	259 (94.5)	74 (16.9)	<0.000***
Crazy paving	65 (23.7)	25 (5.7)	<0.000***
Vascular enlargement	163 (59.5)	19 (4.3)	<0.000***
Reversed halo	17 (6.2)	11 (2.5)	0.014***
Consolidation	12 (4.4)	22 (5)	0.857**

Table 1. Continued			
Tree-in bud	12 (4.4)	16 (3.7)	0.632***
Subpleural bands	69 (25.2)	122 (27.9)	0.423***
Traction bronchiectasis	33 (12)	44 (10.1)	0.409***
Solid nodules	34 (12.4)	53 (12.1)	0.912***
Lymphadenopathy	41 (15)	49 (11.2)	0.146***
Pleural effusion	26 (9.5)	50 (11.4)	0.443***
Mucoid impaction	13 (4.7)	15 (3.4)	0.429**
Bronchial wall thickening	36 (13.1)	63 (14.4)	0.632***

Unless otherwise specified, the data are the number of patients with percentages in parentheses.
*Data are mean±standard deviation. Mann-Whitney U test.
**Fischer's Exact test
***Chi-square test
PCR: Polymerase chain reaction; GGO: Ground-glass opacity

Table 2. Inter-observer agreement of the CO-RADS scoring

	PCR +			PCR -			All patients		
	n (%)	**Kappa**	**%95 CI**	**n (%)**	**Kappa**	**%95 CI**	**n (%)**	**Kappa**	**%95 CI**
CO-RADS 1	17 (6.2)	0.783	0.735-0.832	312 (7.4)	0.771	0.733-0.810	329 (46)	0.861	0.831-0.891
CO-RADS 2	12 (4.4)	0.420	0.372-0.469	46 (10.5)	0.532	0.494-0.571	58 (8.2)	0.503	0.473-0.533
CO-RADS 3	37 (13.5)	0.446	0.397-0.494	30 (6.9)	0.370	0.332-0.408	67 (9.4)	0.408	0.378-0.438
CO-RADS 4	29 (10.6)	0.409	0.361-0.457	15 (3.4)	0.537	0.498-0.575	44 (6.2)	0.468	0.438-0.498
CO-RADS 5	179 (65.3)	0.746	0.697-0.794	34 (7.8)	0.846	0.808-0.884	213 (30)	0.863	0.833-0.893
Overall	274	0.596	0.567-0.625	437	0.651	0.628-0.675	711	0.724	0.706-0.742

PCR: Polymerase chain reaction; CI: confidence interval; CO-RADS: COVID-19 reporting and data system
Fleiss kappa test.

4.1. Diagnostic Accuracy of the Reporting System

After that the four radiologists performed their own scoring, a final scoring was reached for all patients by consensus to evaluate the reporting system. The number of patients scored from CO-RADS 1 to 5 is presented in Table 2. Youden's index showed CO-RADS of ≥ 3 as the optimal threshold to distinguish rRT-PCR (+) patients from rRT-PCR (-) patients. Furthermore, 30 of 437 rRT-PCR (-) patients (6.9%) and 37 of 274 rRT-PCR (+) patients (13.5%) were classified as CO-RADS 3. The ratio of CO-RADS 3 was almost 2 times higher in the rRT-PCR positive patient group than in the rRT-PCR negative group. Due to this difference in rates, accepting this group of patients as COVID-19 negative and not taking precautions may increase the transmission rate of the disease. Moreover, waiting for at least two rRT-PCR results leads to a delay in the final decision. Regarding this, patients classified as CO-RADS 3 should be considered positive in pandemic conditions. Consequently, by considering CO-RADS 3, 4, and 5 positive and CO-RADS 1 and 2 negative, the sensitivity, specificity, positive predictive value, and negative predictive value were 81.9%, 89.4%, 75.7%, and 92.5%, respectively, and the accuracy of the scoring system was 84.8%. The results of the ROC analysis confirmed the diagnostic performance of the CO-RADS (P<0.001) with AUC = 0.857 (95% CI: 0.827-0.886) to predict COVID-19 rRT-PCR positivity (Figure 2). The diagnostic performance of each reader is summarized in Table 3.

When rRT-PCR was considered a gold standard test in the diagnosis of COVID-19, false positivity of CO-RADS 5 was obtained at 34.213 (15%) and false negativity of CO-RADS 1 was estimated at 17.329 (5%). Furthermore, initial and repeated rRT-PCR tests were negative in 15 of 44 patients scored as CO-RADS 4 (figures 3a and 3b).).

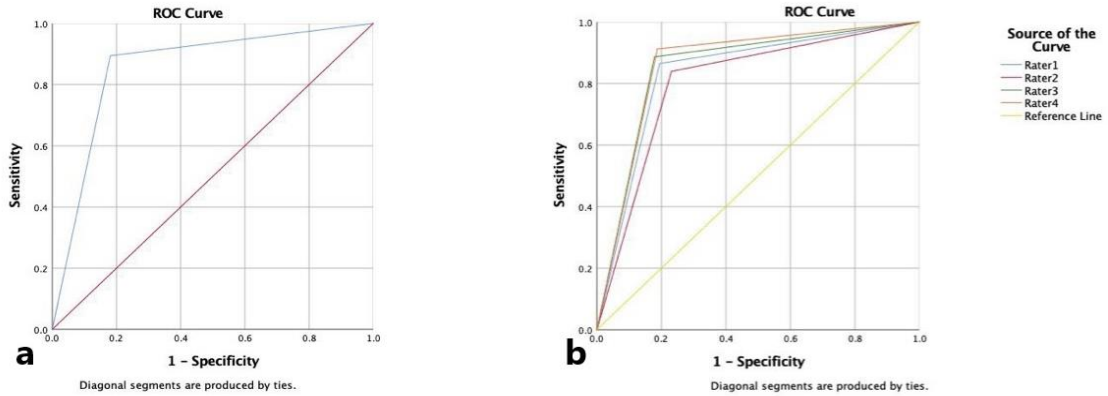

Figure 2. Diagnostic performance of CO-RADS scoring in patients
The area under the receiver operating characteristics curve of CO-RADS for consensus (a) and each rater (b) reflects the prediction of a positive rRT-PCR result in symptomatic individuals

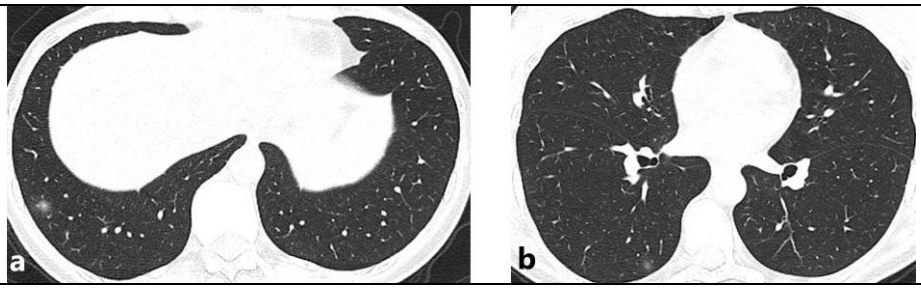

Figure 3. A 39-year-old male patient presented with weakness that had started 4 days ago. He had a negative rRT-PCR test for SARS-CoV-2. Axial non-contrast chest CT scans (a, b) showed ground-glass opacities with a diameter of 8 mm in the lateral-basal segment (a), and a diameter of 5 mm in the postero-basal segment of the right lung (b). The patient was not treated for COVID-19; however, he was quarantined at his home for 14 days due to suspicious CT findings. In the follow-up examination 1 week later, the clinical findings completely regressed, and no additional treatment was planned. In a retrospective evaluation, three radiologists interpreted the CT findings as CO-RADS 4 and one radiologist as CO-RADS 3

Other viral pneumonia agents were the cause in 4 patients, who were scored as CO-RADS 5 and were false positive. Drug intoxication, radiation pneumonitis, and cardiogenic edema were also the causes of false positive results. (figures 4a, 4b, and 4c).

However, 26 of 34 false positive CO-RADS 5 patients might be accepted to be infected by the COVID-19 as the final diagnosis, and at least isolation might be arranged since the rRT-PCR test was not reliable (figures 5a, 5b, and 5c). Considering this, the predicted value of false positivity of CO-RADS 5 was calculated at 8.213 (4%)according to the multidisciplinary approach.

Table 3. Diagnostic performance of the raters with the positive threshold accepted as CO-RADS ≥ 3

	ROC		Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
	AUC	%95 CI				
Rater 1	0.835	0.803-0.867	86.5	80.5	73.6	905
Rater 2	0.804	0.770-0.838	83.9	76.9	69.5	88.4
Rater 3	0.854	0.824-0.884	88.7	82.2	75.7	92.1
Rater 4	0.862	0.833-0.891	91.2	81.2	75.3	93.7
Consensus	0.857	0.827-0.886	81.9	89.4	75.7	92.5

ROC: Receiver operating characteristic; AUC: Area under the curve; CI: Confidence interval; PPV: Positive predictive value; NPV: Negative predictive value

Figure 4. A 34-year-old female patient was transferred from the external center due to drug intoxication (Carbamazepine+Valproic acid+Sertraline). Chest CT was performed on the same day in the patient whose PCR test was negative. Axial CT images (a-c) showed bilateral patchy ground-glass opacities. In addition, bilateral atelectasis (a) and pleural effusion in the left hemithorax (a) were observed. The results of the SARS-CoV-2 PCR test were repeatedly negative on the next day of admission. By analyzing all clinical and laboratory data of the patient, CT findings were accepted as drug-induced pulmonary edema with a multidisciplinary approach. Images were scored as CO-RADS 4 by 3 radiologists and CO-RADS 5 by the other radiologist in the retrospective evaluation and were considered false positive.

Figure 5. A 68-year-old male patient with a history of contact presented with fever and shortness of breath that started 3 days ago. Computed tomography imaging was performed simultaneously with SARS-CoV-2 PCR. Ground glass opacities were observed in axial (a) and coronal (b) section CT images. Vascular enlargement in the ground glass opacity (c) was observed in the superior segment of the right lung lower lobe. Although 2 SARS-CoV-2 PCR tests were negative in this patient, it was evaluated as COVID-19 infection with a multidisciplinary decision due to high suspicion of CT findings. All 4 radiologists scored the CT images as CO-RADS 5.

4.2. Inter-rater Agreement of Grading

Fleiss' kappa was used to quantify the interobserver agreement. The overall κ values and the κ values for CO-RADS 1 to 5 are presented in Table 2. When the researchers evaluated rRT-PCR (-), rRT-PCR (+), and the whole patient group separately, CO-RADS 1 and 5 showed the highest interobserver agreements; however, they were found to be the lowest in CO-RADS 3 and 4. In CO-RADS 1 and 5, observers showed substantial to an almost perfect agreement according to the kappa values, whereas the other values obtained fair to moderate agreement. In addition, the agreement was slightly higher regarding the rRT-PCR (-) patient group than the positive one (Table 3).

5. Discussion

Based on the data published in recent literature, COVID-19 pneumonia had characteristic CT features in the disease process (13-16), such as different degrees of ground-glass opacities with and/or without consolidation, crazy-paving sign, reversed halo sign, thickened vessels within the lesion, multifocal organizing pneumonia, and architectural distortion in a peripheral distribution. In this study, chest CT findings were examined and compared both specifically and non-specifically for COVID-19, in patients with rRT-PCR positive and negative. These two groups were similar in terms of the findings that were not specific for COVID-19 and age distribution. However, we found a significant difference between the two groups regarding the COVID-19-specific chest CT findings and hospitalization (Table 1). Although it is estimated that the hospitalization period due to this infection has increased slightly, part of this difference between the two groups might have been due to our over-treatment. Additionally, it was found that male to female ratio was higher in the rRT-PCR positive patient group than in the rRT-PCR negative one. This result is in line with those of studies stating that severe disease is more common among men (17,18).

Most of the studies evaluating the accuracy and usability of CT imaging, classified chest CT findings as positive or negative for COVID-19 infection (19-24). Nonetheless, in the present study, CT findings were classified into five grades by the consensus of four radiologists since the researchers assumed that it was more useful to evaluate the accuracy of CT findings by grading. De Smet K. et al. (25) investigated the performance of CT-CO-RADS in diagnosing COVID-19 rRT-PCR positivity among symptomatic and asymptomatic individuals. They reported good diagnostic performance with an AUC of 0.89 (95% CI: 0.87-0.91). This finding was in agreement with those of the current study with similar AUC (0.857) on 711 symptomatic individuals. The values achieved in the present research confirmed the probable benefit of using structured reporting of chest CT data in a pandemic setting.

As was mentioned in this study, several reporting systems are created by different groups for the same purpose. Although the naming or enumeration of the categories in these scoring systems is different, the CT findings and setup used in the classification are essentially equal. Chest CT findings described as CO-RADS 5 are equal to the categories defined as "typical" in RSNA chest CT classification, "classic COVID-19" in BSTI guidance statement, and "COVID-RADS 3". The findings covered by CO-RADS 1 and 2 corresponds to "negative for pneumonia" and "atypical appearance" in RSNA chest CT classification, "non-COVID" in BSTI guidance statement, and "COVID-RADS 0" and "COVID-RADS 1" in COVID-RADS. Finally, the findings included in CO-RADS 3 and 4 have been used to create "indeterminate appearance" in RSNA chest CT classification, "indeterminate" and "probable COVID-19" in BSTI guidance statement, and "COVID-RADS 2A" and "COVID-RADS 2B" categories in COVID-RADS. However, not addressing the diseases (i.e., interstitial pneumonia and emphysema) that may coexist with COVID-19 in the scoring system developed by RSNA and not making detailed categorization of the GGO pattern in CO-RADS can be highlighted as small differences. In addition, CO-RADS 0 and CO-RADS 6 groups are defined in CO-RADS, which are neither equivalent to others nor widely used in practice. Therefore, there were no real differences when using CO-RADS, COVID-RADS, RSNA, and BSTI scoring systems. Nevertheless, in the CO-RADS, similar to the scoring systems defined previously (e.g., Prostate Imaging-RADS and Breast Imaging-RADS), the definition of categories by numbers may be more useful in communicating with clinicians who do not have detailed information about radiological findings.

As in other scoring systems, it can be said that the weakest point of the CO-RADS is CO-RADS 3 and CO-RADS 4 distinction. The reason for this may be attributed to the fact that the findings separating these two categories differ slightly from each other and include more subjective definitions. In the current study, it was observed that the most frequent interpretation differences were related to the placement of the peribronchial ground glass areas. In addition, the presence of other accompanying parenchymal findings was another important factor reducing the agreement in scoring. Based on the results, 6.9% and 13.5% of the rRT-PCR negative and positive patient groups were classified as CO-RADS 3, respectively. Similarly, the ratio of CO-RADS 4 was approximately 3 times higher in the rRT-PCR positive group. Therefore, the researchers assumed that CO-RADS 3 patients should be interpreted as positive for COVID infection in practical use until the opposite is proven by at least 3 rRT-PCR negative tests to avoid delay in the final decision phase.

The developers of CO-RADS assessed the interobserver variability with eight radiologists. Their pilot study was performed on a set of 105 CT scans. They reported moderate to substantial interobserver agreement as the κ values were obtained at 0.58 and 0.68 for CO-RADS 1 and CO-RADS 5, respectively. In the present study, the κ values found for CO-RADS 1 and CO-RADS 5 were higher (0.861 and 0.863, respectively) on a much larger sample (n=711). The reason for this discrepancy may be due to the smaller number of radiologists included in this study and to the fact that the radiologists who participated in our study were working in the same center. Moreover, the radiologists participating in our study had more experience in interpreting COVID-suspicious CT, which might be effective in high interobserver agreement. Similarly, the lowest κ values were observed in CO-RADS 3 and CO-RADS 4 (0.408 and 0.468, respectively). In addition to a pilot study, in the current study, the rRT-PCR positive and rRT-PCR negative groups were examined separately in terms of interobserver agreement and it was higher in the rRT-PCR negative group than the positive one (0.651 and 0.596, respectively). This result was predictable in the PCR negative group for which relatively less pathological CT findings were expected.

One of the limitations of the present study was related to its type, which was retrospective. Secondly, our patient group consisted of patients admitted to our hospital during the peak of the pandemic. However, the specificity of CO-RADS might have been lower at a time when other viral pneumonia agents were prevalent. Thirdly, all of the radiologists participating in the study were working in the same center which might have been effective in high interobserver agreement; therefore, a multicenter study could be more beneficial. Lastly, rRT-PCR was used as the gold standard test despite its misconceptions since using the multidisciplinary decision might introduce an affirmation bias.

6. Conclusion

In conclusion, with a useful scoring system, the CT result can be conveyed in a faster and more understandable way and interpretation differences between the clinician and radiologist can be eliminated. Based on the results of this study, the CO-RADS was successful in distinguishing highly suspicious cases in terms of COVID-19 infection pulmonary involvement. In addition, the interobserver agreement observed in this study was better than that in previous studies. Nonetheless, it is required to perform more multicenter studies that evaluate scoring systems. It should also be noted that this scoring system assesses the suspicion of pulmonary involvement in COVID-19 and does not indicate the severity of lung involvement.

Footnotes

Conflicts of Interest: The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Ethical considerations: This retrospective study was approved by the Institutional Review Board of Demiroglu Bilim University (Approved date/number: 13.10.2020/2020-19-04). Written informed consent was waived by the committee.

Main Points:- With positive chest CT findings, patient isolation can be achieved without waiting for the rRT-PCR result.

- Using a CT scoring system in the diagnosis of COVID-19 infection facilitates communication between doctors.

- CO-RADS is a scoring system with high accuracy and interobserver agreement.

- Keeping the circulation fast in the triage area can reduce the COVID-19 transmission.

References

1. World Health Organization. Coronavirus disease (Covid-19) outbreak situation. Geneva: World Health Organization; 2019.
2. Prokop M, van Everdingen W, van Rees Vellinga T, Quarles van Ufford H, Stöger L, Beenen L, et al. CO-RADS: a categorical ct assessment scheme for patients suspected of having COVID-19-definition and evaluation. *Radiology*. 2020;**296**(2):E97-104. doi: 10.1148/radiol.2020201473. [PubMed: 32339082].
3. Li Y, Yao L, Li J, Song Y, Cai Z, Yang C. Stability issues of RT-PCR testing of SARS-CoV-2 for hospitalized patients clinically diagnosed with COVID-19. *J Med Virol*. 2020;**92**(7):903-8. doi: 10.1002/jmv.25786. [PubMed: 32219885].
4. Wang W, Xu Y, Gao R, Lu R, Han K, Wu G, et al. Detection of SARS-CoV-2 in different types of clinical specimens. *JAMA*. 2020;**323**(18):1843-4. doi: 10.1001/jama.2020.3786. [PubMed: 32159775].
5. Corman VM, Landt O, Kaiser M, Molenkamp R, Meijer A, Chu DK, et al. Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. *Euro Surveill*. 2020;**25**(3):2000045. doi: 10.2807/1560-7917.ES.2020.25.3.2000045. [PubMed: 31992387].
6. Hare S, Rodrigues J, Nair A, Robinson G. Lessons from the frontline of the COVID-19 outbreak. London: BMJ Opinion; 2020.
7. Xie X, Zhong Z, Zhao W, Zheng C, Wang F, Liu J. Chest CT for typical coronavirus disease 2019 (COVID-19) Pneumonia: relationship to negative RT-PCR testing. *Radiology*. 2020; **296**(2):E41-5. doi: 10.1148/radiol.2020200343. [PubMed: 32049601].
8. Chan JF, Yuan S, Kok KH, To KK, Chu H, Yang J, et al. A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster. *Lancet*. 2020;**395**(10223):514-23. doi: 10.1016/S0140-6736(20)30154-9. [PubMed: 31986261].
9. Islam N, Salameh JP, Leeflang MM, Hooft L, McGrath TA, van der Pol CB, et al. Thoracic imaging tests for the diagnosis of COVID-19. *Cochrane Database Syst Rev*. 2020;**11**:CD013639. doi: 10.1002/14651858.CD013639. [PubMed: 33242342].
10. Simpson S, Kay FU, Abbara S, Bhalla S, Chung JH, Chung M, et al. Radiological society of north america expert consensus statement on reporting chest CT findings related to COVID-19. endorsed by the society of thoracic radiology, the American College of

Radiology, and RSNA - Secondary Publication. *J Thorac Imaging*. 2020;**35**(4):219-27. doi: 10.1097/RTI.0000000000000524. [PubMed: 32324653].
11. Salehi S, Abedi A, Balakrishnan S, Gholamrezanezhad A. Coronavirus disease 2019 (COVID-19) imaging reporting and data system (COVID-RADS) and common lexicon: a proposal based on the imaging data of 37 studies. *Eur Radiol*. 2020;**30**(9):4930-42. doi: 10.1007/s00330-020-06863-0. [PubMed: 32346790].
12. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;**33**(1):159-74. doi: 10.2307/2529310.
13. Rubin EJ, Baden LR, Morrissey S, Campion EW. Medical Journals and the 2019-nCoV Outbreak. *N Engl J Med*. 2020;**382**(9):866. doi: 10.1056/NEJMe2001329. [PubMed: 31986242].
14. Gralinski LE, Menachery VD. Return of the coronavirus: 2019-nCoV. *Viruses*. 2020;**12**(2):135. doi: 10.3390/v12020135. [PubMed: 31991541].
15. Duarte R, Furtado I, Sousa L, Carvalho CFA. The 2019 novel coronavirus (2019-nCoV): novel virus, old challenges. *Acta Med Port*. 2020;**33**(3):157-7. doi: 10.20344/amp.13547. [PubMed: 32023427].
16. Wang M, Cao R, Zhang L, Yang X, Liu J, Xu M, et al. Remdesivir and chloroquine effectively inhibit the recently emerged novel coronavirus (2019-nCoV) in vitro. *Cell Res*. 2020;**30**(3):269-71. doi: 10.1038/s41422-020-0282-0. [PubMed: 32020029].
17. Sharma G, Volgman AS, Michos ED. Sex differences in mortality from COVID-19 pandemic: are men vulnerable and women protected? *JACC Case Rep*. 2020;**2**(9):1407-10. doi: 10.1016/j.jaccas.2020.04.027. [PubMed: 32373791].
18. Klein SL, Dhakal S, Ursin RL, Deshpande S, Sandberg K, Mauvais-Jarvis F. Biological sex impacts COVID-19 outcomes. *PLoS Pathog*. 2020;**16**(6):e1008570. doi: 10.1371/journal.ppat.1008570. [PubMed: 32569293].
19. Huang P, Liu T, Huang L, Liu H, Lei M, Xu W, et al. Use of chest CT in combination with negative RT-PCR assay for the 2019 novel coronavirus but high clinical suspicion. *Radiology*. 2020; **295**(1):22-3. doi: 10.1148/radiol.2020200330. [PubMed: 32049600].
20. Lei J, Li J, Li X, Qi X. CT imaging of the 2019 novel coronavirus (2019-nCoV) pneumonia. *Radiology*. 2020;**295**(1):18. doi: 10.1148/radiol.2020200236. [PubMed: 32003646].
21. Shi H, Han X, Zheng C. Evolution of CT manifestations in a patient recovered from 2019 novel coronavirus (2019-nCoV) pneumonia in Wuhan, China. *Radiology*. 2020;**295**(1):20. doi: 10.1148/radiol.2020200269. [PubMed: 32032497].
22. Ai T, Yang Z, Hou H, Zhan C, Chen C, Lv W, et al. 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. doi: 10.1148/radiol.2020200642. [PubMed: 32101510].
23. Fang Y, Zhang H, Xie J, Lin M, Ying L, Pang P, et al. Sensitivity of chest CT for COVID-19: comparison to RT-PCR. *Radiology*. 2020;**296**(2):E115-7. doi: 10.1148/radiol.2020200432. [PubMed: 32073353].
24. Long C, Xu H, Shen Q, Zhang X, Fan B, Wang C, et al. Diagnosis of the coronavirus disease (COVID-19): rRT-PCR or CT? *Eur J Radiol*. 2020;**126**:108961. doi: 10.1016/j.ejrad.2020.108961. [PubMed: 32229322].
25. De Smet K, De Smet D, Ryckaert T, Laridon E, Heremans B, Vandenbulcke R, et al. Diagnostic performance of chest CT for SARS-CoV-2 infection in individuals with or without COVID-19 symptoms. *Radiology*. 2021; **298**(1):E30-7. doi: 10.1148/radiol.2020202708. [PubMed: 32776832].