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

COVID-19 Enfeksiyonu Şüphesi ile Başvuran Hastalarda Akciğer Tomografi Bulgularının Raporlama Sistemi Dahilinde Değerlendirilmesi

Ibrahim Halil Sever1, Bahattin Özkul2, Kourosh Yaghouti3, Bedriye Koyuncu Sökmen1, Nagihan İnan Gürcan1

1Demiroğlu Bilim University, Florence Nightingale Hospital, Clinic of Radiology, İstanbul, Turkey
2Atlas University Medicine Hospital, Clinic of Radiology, İstanbul, Turkey

Abstract

Objective: To evaluate the diagnostic performance of Coronavirus disease-2019 (COVID-19) reporting and data system (CO-RADS) in patients admitted with suspected COVID-19 infection.

Method: This retrospective study included all patients admitted to our hospital with COVID-19 pneumonia suspicion between March 20 and May 15 2020, who were examined by both computed tomography (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 patients’ medical history, assessed all images independently and classified the CT findings according to the CO-RADS that was 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 the present study, 274 [153 men, 121 women, median age=49, interquartile range (IQR): 25-62 years] rRT-PCR positive and 437 (208 men, 229 women; median age=46; IQR 32-64 years) rRT-PCR negative individuals were included. CO-RADS had a good diagnostic performance with area under the curve of 0.857. The sensitivity, specificity, positive predictive value, negative predictive value and accuracy were 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 fair to moderate agreement. The intrarater agreement in the PCR (-) patient group was slightly higher than the positive one.

 Öz

Amaç: Koronavirüs hastalığı-2019 (COVID-19) enfeksiyonu şüphesi bulunan hastalarda CO-RADS’ı tansal performansını değerlendiririz.

Yöntem: Retrospektif olarak gerçekleştirildiğimiz çalışmamız, 20 Mart-15 Mayıs 2020 tarihleri arasında COVID-19 pnömoni şüphesiyle hastaneye başvuran, ilk başvuruyuzda hem bilgisayarlı tomografi hem de gerçek zamanlı ters transkripsiyon-polimeraz zincir reaksiyonu (rRT-PCR) ile tetkik edilen tüm hastalar dahil edildi. Hastaların laboratuvur sonuçları ve tıbbi geçmişleri hakkında bilgisiz olan dört radyolog, CO-RADS sisteminin tansal değerini ve tüm CO-RADS sınıfları için verilen değerlere göre özelliklerini araştırarak değerlendirildi. Raporlama sisteminin tansal değeri ve tüm CO-RADS sınıfları için verilen değerlere göre özelliklerini araştırarak değerlendirildi.

Bulgular: Çalışma grubumuz 274 [153 erkek, 121 kadın; Orta yaş=49, çeyrek interquartile aralığı (ÇAA): 25-62 yaş] rRT-PCR pozitifi ve 437 (208 erkek, 229 kadın; Orta yaş=46, ÇAA: 32-64 yaş) rRT-PCR negatif hastanın oluşma eğrini, CO-RADS, 0.857 eğrinin altındaki alanı değerine ve yüksek bir tansal performansı sahipti. Duyarlılık, özgülük, pozitif öngörü değeri, negatif öngörü değeri ve doğruluk sırasıyla %81,9, %89,4, %75,7, %92,5 ve %84,8 idi. CO-RADS 1 ve 5te dört radyoloğun ölçümü, kapa degerlerine göre sırasıyla 0,60-0,80 ve 0,80-1 aralığında, diğer CO-RADS sınıflarına ise 0,20-0,40 ve 0,40-0,60 arasında hesaplandı. Son olarak PCR negatif hasta grubundaki özellikler arasındaki uyum, PCR pozitifi olanlarda daha yüksek idi.
Introduction

In December 2019, an outbreak with respiratory symptoms caused by a novel-coronavirus-2 (SARS-CoV-2) appeared in Wuhan, China and due to the high rate of human-to-human transmission, it spread almost all countries in a short while and World Health Organization declared it as pandemic on March 11, 2020. (1) And now, at the end of 2020, about 84 million people were infected and almost 1.8 million people died 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 to 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 very recent, CT findings described in the literature are open to interpretation differences 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. It is obvious that clinicians and radiologists need a clearer statement about what the observed findings mean in order 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 among physicians and to 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 a 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: Negative for pneumonia, atypical appearance, indeterminate appearance and typical appearance (10). Another group of researchers called their systems as 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 COVID-19 reporting and data system (CO-RADS) that correlated with the scoring systems previously defined and classified chest CT findings according to scores ranging from 0 to 6 (2).

In this study, we aimed to evaluate the diagnostic performance of CO-RADS not only in rRT-PCR positive patients but also in all patients who applied with the suspicion of COVID-19 infection and underwent chest CT.

Materials and Methods

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

Patients

From March 20 to May 15, 2020, 711 patients (274 rRT-PCR positive and 437 rRT-PCR negative) who underwent non-contrast chest tomography due to clinical suspicion.
of COVID-19 infection (ie, fever higher > 37.5 °C, cough, and/or shortness of breath) were enrolled in this study. Nasopharyngeal and oropharyngeal swab specimens were taken from all patients by using COVID-19 specific kits (Bio-Speedy® SARS-CoV-2 triple gene RT-qPCR, Bioeksen, Istanbul, Turkey). Patients who showed at least one positive RT-PCR were considered to be positive for COVID-19; otherwise, if CT findings were suggestive of COVID-19, repeated RT-PCR testing up to a maximum of three times in 14 days were performed. Patients with initial negative RT-PCR and negative CT findings underwent a 14-day follow-up. These patients were considered negative if any symptoms consistent with COVID-19 worsened or laboratory findings did not occur during follow-up. All CT scans were performed within the first 3 days after blood samples were taken. Patients whose CT images could not be evaluated clearly due to artefacts were excluded from the study (CO-RADS 0). Furthermore, we did not use CO-RADS 6 as the evaluation of the images was done completely without knowledge of the lab result.

Image Analysis
Four radiologists (with 10, 12, 12 and 18 years of experience in chest imaging, respectively), who were blinded to the rRT-PCR results and patients’ medical history, assessed all images independently by using 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 7 different groups. 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 very low level of suspicion (Figure 1b), CO-RADS 2 as low level of suspicion (Figure 1c), CO-RADS 3 as equivocal findings (Figure 1d), CO-RADS 4 as high level of suspicion (Figure 1e), and CO-RADS 5 as 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.

Statistical Analysis
All statistical analyses were performed by IBM Statistical Package for the Social Sciences (SPSS version 25 for macOS, Chicago, IL, USA). The fitness of numeric data set to normal distribution was determined by the Kolmogorov-Smirnov test, and for the non-normal distribution, median and interquartile range (IQR) were used to express measurement data. CT findings of patients between rRT-PCR positive and negative groups were compared with $\chi^2$ 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 interobserver agreement and to determine the level of agreement between all observers. κ-values were divided into groups as described by Landis and Koch: 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) (12). 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. Also, for each radiologist, the highest Youden index (J=sensitivity + specificity-1) was calculated for each reader to select the optimal threshold to discriminate between COVID-19 positive and negative patients, and the corresponding sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) were calculated. p<0.05 was considered statistically significant.

Results
In our retrospective cohort of 753 patients, 711 individuals were included in the study. Two hundred seventy-four patients (153 men, 121 women; median age=49; IQR 25-62 years) were rRT-PCR positive and 437 patients (208 men, 229 women; median age=46; IQR 32-64 years) were rRT-PCR negative. 28 patients were excluded from the study when the time-interval between chest CT and the rRT-PCR assay was longer than 3 days. Also, 14 patients were excluded due to 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 patients had two and 24 patients had three tests. 176 of 437 rRT-PCR negative patients had two tests and others had one.

We compared PCR positive and PCR negative groups 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).

**Diagnostic Accuracy of the Reporting System**

After four radiologists performed their own scoring, a final scoring was reached for all patients by consensus in order to evaluate the reporting system. Number of patients scored from CO-RADS 1 to 5 are supplied in Table 2.

Youden’s index ($J=0.713$) showed CO RADS ≥3 as the optimal threshold to distinguish rRT-PCR (+) patients from rRT-PCR (-) patients. 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 in the rRT-PCR positive patient group was almost 2 times higher 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 COVID-19.

### Table 1. Demographics and chest CT findings of patients

| Characteristics          | PCR +   | PCR -   | $p$ (2-tailed) |
|--------------------------|---------|---------|----------------|
| Age (years)*             | 49 (25-62) | 46 (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         |
| Reverse halo             | 17 (6.2)  | 11 (2.5)  | 0.014          |
| Consolidation            | 12 (4.4)  | 22 (5)    | 0.857**        |
| 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, data are number of patients, with percentages in parentheses. *Data are median, with ranges in parentheses, **Fisher’s Exact test, PCR: Polymerase chain reaction, GGO: Ground glass opacity
the disease. Moreover, waiting at least two rRT-PCR results leads to a delay in the final decision. That’s why, patients classified as CO-RADS 3 should be considered positive in pandemic conditions. So, when we considered CO-RADS 3, 4 and 5 positive, CO-RADS 1 and 2 negative, the sensitivity, specificity, PPV and NPV were 81.9%, 89.4%, 75.7% and 92.5%, respectively, and the accuracy of the scoring system was 84.8%. ROC analysis confirmed the diagnostic performance (p<0.001) of CO-RADS 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 shown in Table 3.

When we consider rRT-PCR as gold standard test in the diagnosis of COVID-19, false positivity of CO-RADS 5 was 34/213 (15%) and false negativity of CO-RADS 1 was 17/329 (5%). Moreover, initial and repeated rRT-PCR tests were negative in 15 of 44 patients scored as CO-RADS 4 (Figure 3a, b). Other viral panels (Sofia®, Quidel, San Diego, USA) were applied in 32 out of 430 patients in total, and the cause was other viral pneumonia agents in 4 patients scored as CO-RADS 5 and were false positive. Drug intoxication, radiation pneumonitis and cardiogenic edema were also the causes of false positive results (Figure 4a-c). However, 26 of 34 false positive CO-RADS 5 patients might be accepted to be COVID-19 infection as final diagnosis and at least isolation might be arranged since the rRT-PCR test is not reliable (Figure 5a-c). That’s why, the predicted value of false positivity of CO-RADS 5 was 8/213 (4%) according to the multidisciplinary approach.

| 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 (71.4) | 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 | - | 0.596 | 0.567-0.625 | - | 0.651 | 0.628-0.675 | - | 0.724 | 0.706-0.742 |

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

| **ROC AUC** | **95% CI** | **Sensitivity (%)** | **Specificity (%)** | **PPV (%)** | **NPV (%)** |
|-------------|------------|---------------------|---------------------|-------------|-------------|
| Rater 1     | 0.835      | 0.803-0.867         | 86.5                | 80.5        | 73.6        | 90.5        |
| 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        |

**PCR:** Polymerase chain reaction, **CI:** Confidence interval, **n:** Number of patients, **CO-RADS:** COVID-19 reporting and data system, **COVID-19:** Coronavirus disease-2019

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

CO-RADS: Coronavirus disease-2019 reporting and data system, rRT-PCR: Real-time reverse transcription-polymerase chain reaction
Interreader Agreement of Grading

Fleiss’ kappa was used to quantify the interobserver agreement. The overall kappa values and the kappa values for CO-RADS 1 to 5 were shown in Table 2. When we evaluated rRT-PCR (−), rRT-PCR (+) and the whole patient group separately, CO-RADS 1 and 5 showed the highest interobserver agreements; however, they were lower in CO-RADS 3 and 4. In CO-RADS 1 and 5, observers showed substantial to almost perfect agreement according to the kappa values whereas others showed fair to moderate agreement. In addition, the agreement in the rRT-PCR (−) patient group was slightly higher than the positive one (Table 3).

Discussion

On the basis of 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, revers halo sign, thickened vessels within the lesion, multifocal organizing pneumonia and architectural distortion in a peripheral distribution. In this study, we examined and compared chest CT findings, both specific and non-specific 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 two groups in terms COVID-19 specific chest CT findings and hospitalization (Table 1). Although it is estimated that hospitalization period due to this infection has increased slightly, part of this difference between the two groups may have been due to our over-treatment. Also, we found that male to female ratio in rRT-PCR positive patient group was higher than in the rRT-PCR negative group. This result supported the articles in the literature stating that severe disease was more common in men (17,18).

Figure 3a, b. Thirty-nine-year-old male patient presented with weakness that started 4 days ago. He had a negative rRT-PCR test for SARS-CoV-2. Axial non-contrast chest CT scans (a,b) show ground glass opacities with a diameter of 8 mm in the lateral-basal segment (a), and with a diameter of 5 mm in the postero-basal segment of the right lung (b). The patient was not treated for COVID-19, but 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 retrospective evaluation, three radiologists interpreted the CT findings as CORADS-4 and one radiologist as CO-RADS 3.

COVID-19: Coronavirus disease-2019, CT: Computed tomography, SARS-CoV-2: Severe acute respiratory syndrome-coronavirus-2, CO-RADS: Coronavirus disease-2019 reporting and data system

Figure 4a-c. Thirty-four-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) show bilateral patchy ground glass opacities. In addition, bilateral atelectasis (a) and pleural effusion in the left hemithorax (a) are seen. SARS-CoV-2 PCR test was repeatedly negative in 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 as CO-RADS 5 by the other radiologist in retrospective evaluation and were considered false positive.

COVID-19: Coronavirus disease-2019, CT: Computed tomography, SARS-CoV-2: Severe acute respiratory syndrome-coronavirus-2, CO-RADS: Coronavirus disease-2019 reporting and data system, PCR: Polymerase chain reaction

Figure 5a-c. Sixty-eight-year-old male patient with a history of contact presented with fever and shortness of breath that started 3 days ago. CT imaging was performed simultaneously with SARS-CoV-2 PCR. Ground glass opacities are seen in axial (a) and coronal (b) section CT images. Vascular enlargement in the ground glass opacity (c) is observed in the superior segment of 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.

COVID-19: Coronavirus disease-2019, CT: Computed tomography, SARS-CoV-2: Severe acute respiratory syndrome-coronavirus-2, CO-RADS: Coronavirus disease-2019 reporting and data system, PCR: Polymerase chain reaction
Indeed, 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). Whereas we classified CT findings into five grades by consensus of four radiologists because we think that it is more useful to evaluate the accuracy of CT findings by grading. De Smet et al. (25) investigated the performance of CT-CORADS to diagnose COVID-19 rRT-PCR positivity in symptomatic and asymptomatic individuals. They reported a good diagnostic performance with AUC under the ROC curve of 0.89 (95% CI: 0.87-0.91). Our study supports the result of this study with similar AUC (0.857) on 711 symptomatic individuals. The values we have achieved confirms that using of structured reporting of chest CT data may be useful in pandemic setting.

As we mentioned, there are several reporting systems created by different groups for the same purpose. Although the naming or enumeration of the categories in these scoring systems are 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”. Findings covered by CORADS 1 and 2 are similar 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 (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 not equivalent to others and are not widely used in practice. Therefore, not surprisingly, there were no real differences when using CO-RADS, COVID-RADS, RSNA and BSTI scoring systems. Yet, in the CO-RADS, similar to the scoring systems defined previously (ie; PI-RADS, BI-RADS etc.), 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, we can say that the weakest point of this system is CO-RADS 3 and CO-RADS 4 distinction. The reason for this may be that the findings that separate these two categories differ slightly and include more subjective definitions. We 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. 6.9% of the rRT-PCR negative patient group was classified as CO-RADS 3. This ratio was 13.5% in the rRT-PCR positive patients. Similarly, the ratio of CO-RADS 4 was approximately 3 times higher in the rRT-PCR positive group. Therefore, we think 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 in order to avoid delay in the final decision phase.

The developers of CO-RADS assessed the interobserver variability with 8 radiologists. Their pilot study was performed on a set of 105 CT scans. They reported moderate to substantial interobserver agreement as the Fleiss’ kappa values were 0.58 for CO-RADS 1 and 0.68 for CO-RADS 5. The kappa values we found for CO-RADS 1 and CO-RADS 5 were higher (0.861 and 0.863, respectively) on a much larger sample (711 patients). This may be because of that less number of radiologists were included in the study and 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 also may be effective in high interobserver agreement. Similarly, lower kappa values were observed in CO-RADS 3 and CO-RADS 4 (0.408 and 0.468, respectively). In addition to pilot study, we examined the rRT-PCR positive and rRT-PCR negative groups separately in terms of interobserver agreement and it was higher in 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 we expected relatively less pathological CT findings.

**Study Limitations**

Our study has some limitations. First, this was a retrospective study. Second, our patient group consisted of patients who applied to our hospital during the peak of the pandemic. However, the specificity of CO-RADS may have been lower at a time when other viral pneumonia agents were prevalent. Third, all of the radiologists participating in the study were working in the same center, which may have been effective in high interobserver agreement, therefore...
a multicenter study could be more beneficial. Fourth, all serological tests of rRT-PCR negative patients could not be performed on the day of the study. Last, we used rRT-PCR as the gold standard test despite its misconceptions because using the multidisciplinary decision may introduce an affirmation bias.

**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. The results we obtained in this study show that the CO-RADS is successful in distinguishing highly suspicious cases in terms of COVID-19 infection pulmonary involvement. In addition, the interobserver agreement we observed was better than in previous studies. However, we think that more multicenter studies that evaluate scoring systems are needed and it should be noted that this scoring system assesses the suspicion of pulmonary involvement in COVID-19 and does not indicate the severity of lung involvement.

**Ethics**

**Ethics Committee Approval:** This retrospective study was approved by the Institutional Review Board of Demiroglu Bilim University (approved date/number: 13.10.2020/2020-19-04).

**Informed Consent:** Written informed consent was waived by the committee.

**Peer-review:** Externally peer-reviewed.

**Authorship Contributions**

Concept: I.H.S., B.Ö., N.I.G., Design: K.Y., B.K.S., N.I.G., Data Collection or Processing: I.H.S., B.Ö., K.Y., Analysis or Interpretation: K.Y., B.K.S., Literature Search: B.Ö., B.K.S., Writing: I.H.S., Manuscript Review and Revisation: N.I.G.

**Conflict of Interest:** No conflict of interest was declared by the authors.

**Financial Disclosure:** The authors declared that this study received no financial support.

**References**

1. World Health Organization, Coronavirus disease (Covid-19) outbreak situation. Available from: https://www.who.int/emergencies/diseases/novel-coronavirus-2019.

2. Prokop M, van Everdingen W, van Rees Vellinga T, Quarles van Ufford H, Stöger L, Beenen L, et al; COVID-19 Standardized Reporting Working Group of the Dutch Radiological Society. CO-RADS: a categorical CT assessment scheme for patients suspected of having COVID-19-definition and evaluation. Radiology 2020;296(2):E97-E104. doi: 10.1148/radiol.2020201473.

3. Li Y, Yao L, Li J, Chen L, 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):903-908.

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):1843-1844.

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.

6. Hare S, Rodrigues J, Nair A, Robinson G. Lessons from the frontline of the covid-19 outbreak. In: BMJ Opin https://blogs.bmj.com/bmj/2020/03/20/lessons-from-the-frontline-of-the-covid-19-outbreak/.

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-E45. doi: 10.1148/radiol.2020200343.

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):514-523.

9. Thoracic Imaging in COVID-19 Infection, Guidance for the Reporting Radiologist (British Society of Thoracic Imaging). Last Accessed Date: 17.10.2020. Available from: https://www.bsti.org.uk/media/resources/files/BSTI_COVID-19_Radiology_Guidance_version_2_16.03.20.pdf.

10. Simpson S, Kay FU, Abbara S, Bhalla S, Chung JH, Chung M, et al. Radiological Society of North America Expert Consensus Document on reporting chest CT findings related to COVID-19: endorsed by the Society of Thoracic Radiology, the American College of Radiology, and RSNA. Radiol Cardiothorac Imaging 2020;2(2):e200152. doi: 10.1148/ryct.2020200152.

11. Salehi S, Abedi A, Balakrishnan S, Gholamrezaezhad 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):4930-4942.

12. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33(1):159-174.

13. Rubin EJ, Baden LR, Morrissey S, Campion EW. Medical journals and the 2019-nCoV outbreak. N Engl J Med 2020;382(9):866.

14. Gralinski LE, Menachery VD. Return of the coronavirus: 2019-nCoV, viruses 2020;24;12(2):135.

15. Duarte R, Furtado I, Sousa L, Carvalho CFA. The 2019 novel coronavirus (2019-nCoV): novel virus, old challenges. Acta Med Port 2020 Feb 5. doi: 10.20344/amp.13547. Epub ahead of print.

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):269-271.

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):1407-1410.
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.

19. Huang P, Liu T, Huang L, Liu H, Lei M, Xu W. Use of chest CT in combination with negative rt-PCR assay for the 2019 novel coronavirus but high clinical suspicion. Radiology 2020;295(1):22-23.

20. Lei J, Li J, Li X, Qi X. CT Imaging of the 2019 Novel Coronavirus (2019-nCoV) Pneumonia. Radiology 2020;295(1):18.

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.

22. 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-E40. doi: 10.1148/radiol.2020200642.

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-E117. doi: 10.1148/radiol.2020200432.

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.

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-E37. doi: 10.1148/radiol.2020202708.