CO-RADS: a categorical CT assessment scheme for patients suspected of having COVID-19 — definition and evaluation

Prokop, M.; Everdingen, W. van; Vellinga, T.V.; Ufford, H.Q. van; Stoger, L.; Beenen, L.; ... ; Dutch Radiological Soc

Radiology

ORIGINAL RESEARCH • THORACIC IMAGING

CO-RADS: A Categorical CT Assessment Scheme for Patients Suspected of Having COVID-19—Definition and Evaluation

Mathias Prokop, MD, PhD • Wouter van Everdingen, MD, PhD* • Tjalco van Rees Vellinga, MD • Henriëtte Quarles van Ufford, MD, PhD • Lauran Stöger, MD, PhD • Ludo Beenen, MD • Bram Geurts, MD • Hester Gietema MD, PhD • Jasenko Krdzalic, MD • Cornelia Schaefer-Prokop, MD, PhD • Bram van Ginneken MD, PhD • Monique Brink, MD, PhD • for the COVID-19 Standardized Reporting Working Group of the Dutch Radiological Society*

From the Department of Radiology, Nuclear Medicine and Anatomy Radboudumc, Radboud University Medical Center, PO Box 9101, 6500 HB Nijmegen, the Netherlands (M.P., W.v.E., B.G., B.v.G., M.B.); Department of Radiology and Nuclear Medicine, Haaglanden Medical Center, The Hague, the Netherlands (T.v.R.V., H.Q.v.U.); Department of Radiology, Leiden University Medical Centre, Leiden, the Netherlands (L.S.); Department of Radiology and Nuclear Medicine, Amsterdam UMC, Amsterdam, the Netherlands (L.B.); Department of Radiology and Nuclear Medicine, Maastricht University Medical Center+, the Netherlands (H.G.); GROW School of Oncology and Developmental Biology, Maastricht, the Netherlands (H.G.); Department of Radiology, Zuyderland MC, Heerlen, the Netherlands (J.K.); and Department of Radiology, Meander Medical Center, Amersfoort, the Netherlands (C.S.). Members of the COVID-19 Standardized Reporting Working Group of the Dutch Radiological Society are listed in Appendix E4 (online). Received April 8, 2020; revision requested April 14; revision received April 17; accepted April 23. **Address correspondence to** M.B. (e-mail: *monique.brink@radboudumc.nl*).

* M.P. and W.v.E. contributed equally to this work.

Conflicts of interest are listed at the end of this article.

Radiology 2020; 296:E97–E104 • https://doi.org/10.1148/radiol.2020201473 • Content code: CH

Background: A categorical CT assessment scheme for suspicion of pulmonary involvement of coronavirus disease 2019 (COVID-19 provides a basis for gathering scientific evidence and improved communication with referring physicians.

Purpose: To introduce the COVID-19 Reporting and Data System (CO-RADS) for use in the standardized assessment of pulmonary involvement of COVID-19 on unenhanced chest CT images and to report its initial interobserver agreement and performance.

Materials and Methods: The Dutch Radiological Society developed CO-RADS based on other efforts for standardization, such as the Lung Imaging Reporting and Data System or Breast Imaging Reporting and Data System. CO-RADS assesses the suspicion for pulmonary involvement of COVID-19 on a scale from 1 (very low) to 5 (very high). The system is meant to be used in patients with moderate to severe symptoms of COVID-19. The system was evaluated by using 105 chest CT scans of patients admitted to the hospital with clinical suspicion of COVID-19 and in whom reverse transcription-polymerase chain reaction (RT-PCR) was performed (mean, 62 years ± 16 [standard deviation]; 61 men, 53 with positive RT-PCR results). Eight observers used CO-RADS to assess the scans. Fleiss κ value was calculated, and scores of individual observers were compared with the median of the remaining seven observers. The resulting area under the receiver operating characteristics curve (AUC) was compared with results from RT-PCR and clinical diagnosis of COVID-19.

Results: There was absolute agreement among observers in 573 (68.2%) of 840 observations. Fleiss κ value was 0.47 (95% confidence interval [CI]: 0.45, 0.47), with the highest κ value for CO-RADS categories 1 (0.58, 95% CI: 0.54, 0.62) and 5 (0.68, 95% CI: 0.65, 0.72). The average AUC was 0.91 (95% CI: 0.85, 0.97) for predicting RT-PCR outcome and 0.95 (95% CI: 0.91, 0.99) for clinical diagnosis. The false-negative rate for CO-RADS 1 was nine of 161 cases (5.6%; 95% CI: 1.0%, 10%), and the false-positive rate for CO-RADS category 5 was one of 286 (0.3%; 95% CI: 0%, 1.0%).

Conclusion: The coronavirus disease 2019 (COVID-19) Reporting and Data System (CO-RADS) is a categorical assessment scheme for pulmonary involvement of COVID-19 at unenhanced chest CT that performs very well in predicting COVID-19 in patients with moderate to severe symptoms and has substantial interobserver agreement, especially for categories 1 and 5.

Online supplemental material is available for this article.

Definitive diagnosis of coronavirus disease 2019 (COVID-19) is usually made by using a reverse transcription-polymerase chain reaction (RT-PCR) assay, which performs accurately in a laboratory setting. However, reported sensitivities in clinical practice range from 42% to 83% and depend on symptom duration, viral load, and test sample quality (1–5). Cases are increasingly reported in which the assay yielded a positive result only after multiple negative results in patients with typical clinical and imaging signs of COVID-19 (6,7). Also, RT-PCR takes hours, or even days, before the results are available, putting strain on the holding units where patients are kept before being sent to a normal or COVID-19 ward. Increasingly, situations arise in which RT-PCR tests are scarce and cannot be used for every patient.

In light of this, the role of chest CT in patients suspected of having COVID-19 is constantly evolving with modest scientific evidence but substantial differences in opinion on when and how the technique should be used for clinical work-up or treatment decisions. Although the seventh edition of the Chinese Novel Coronavirus Pneumonia Diagnosis and Treatment Plan incorporates CT

Abbreviations

AUC = area under the receiver operating characteristics curve, BI-RADS = Breast Imaging Reporting and Data System, CI = confidence interval, CO-RADS = COVID-19 Reporting and Data System, COVID-19 = coronavirus disease 2019, RSNA = Radiological Society of North America, RT-PCR = reverse transcription-polymerase chain reaction, SARS-CoV-2 = severe acute respiratory syndrome coronavirus 2

Summary

The Coronavirus disease 2019 (COVID-19) Reporting and Data System is a categorical assessment scheme for chest CT in patients suspected of having COVID-19, representing the level of suspicion for pulmonary involvement; the substantial agreement among observers and its discriminatory value make it well-suited for use in clinical practice.

Key Results

- Coronavirus disease 2019 (COVID-19) Reporting and Data System (CO-RADS) provided a standardized assessment scheme that simplifies reporting with a five-point scale of suspicion for pulmonary involvement of COVID-19 at chest CT.
- CO-RADS had moderate to substantial agreement among observers, with an overall Fleiss κ value of 0.47 (95% confidence interval [CI]: 0.45, 0.49).
- The discriminatory power of CO-RADS in the diagnosis of COVID-19 was high, with a mean area under the receiver operating characteristic curve of 0.91 (95% CI: 0.85, 0.97) for positive reverse transcription–polymerase chain reaction results.

imaging into the criteria that clinically define COVID-19 (8), the American College of Radiology, among others, discouraged the use of CT in the initial work-up of patients (9,10), only advocating its use for problem solving. The Fleischner Society, in a recent statement, however, sees a role for imaging in various scenarios, with imaging, in particular CT scanning, as a major tool to use if symptoms worsen or in an environment that is resource-constrained for RT-PCR (11).

COVID-19 has CT findings that partially overlap with those of other diseases, mainly viral infections, but it also has characteristic features that are seen less frequently in other settings (12). Various attempts have been made to standardize CT reporting in patients suspected of having COVID-19. The recent Radiological Society of North America (RSNA) Expert Consensus Statement on Reporting (13), for example, proposes standardized nomenclature and an imaging classification for COVID-19 pneumonia that involves four categories (ie, typical appearance, indeterminate appearance, atypical appearance, and negative for pneumonia).

In early March 2020, the Dutch Radiological Society (Nederlandse Vereniging voor Radiologie) initiated a COVID-19 network to facilitate development and nationwide dissemination of COVID-19–related information and tools. Within this network, a COVID-19 standardized reporting working group was formed. The authors developed a standardized assessment scheme for pulmonary involvement of COVID-19 that would make it possible to compare data across institutions and populations and, thus, provide a basis for gathering scientific evidence and improved communication with referring physicians.

Because the system is based on other efforts for standardization, such as Lung Imaging Reporting and Data System (hereafter, LI-RADS), Prostate Imaging Reporting and Data System (hereafter, PI-RADS), or Breast Imaging Reporting and Data System (hereafter, BI-RADS), the authors chose the term COVID-19 Reporting and Data System (CO-RADS) (14). The system was iteratively refined through feedback from members and input from clinical partners. This type of system has been shown to work well in clinical practice and to allow for selection of optimal cutoff points for various clinical decisions depending on the tasks at hand. The current version represents the consensus formed on April 7, 2020.

In this article, the definitions of CO-RADS are presented, along with the results of an initial observer study to assess the interobserver variability and diagnostic accuracy of the proposed system in the hands of observers with variable experience in reading chest CT scans for suspected COVID-19.

CO-RADS, the COVID-19 Reporting and Data System

CO-RADS provides a level of suspicion for pulmonary involvement of COVID-19 based on the features seen at unenhanced chest CT. The level of suspicion increases from very low (CO-RADS category 1) to very high (CO-RADS category 5). Two additional categories encode a technically insufficient examination (CO-RADS category 0) and RT-PCR–proven severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection at the time of examination (CO-RADS category 6).

It should be noted that CO-RADS is a CT-based system that is used to assess the suspicion of pulmonary involvement in COVID-19. The actual interpretation of whether a patient has COVID-19 needs to include other data, such as laboratory test results, clinical findings, and type and duration of symptoms. At present, the reference standard for diagnosing COVID-19 remains positive RT-PCR results. In clinical practice, however, this may require repeated testing that includes deep bronchial and fecal samples and may be hampered by scarcity of tests in high-prevalence areas.

An overview of CO-RADS is given in Table 1, and a pictorial overview is presented in Appendix E1 (online).

CO-RADS Category 0

CO-RADS category 0 is chosen if none of the five categories can be assigned because scans are incomplete or of insufficient quality, for example because of severe artifacts due to coughing or breathing.

CO-RADS Category 1

CO-RADS category 1 implies a very low level of suspicion for pulmonary involvement by COVID-19 based on either normal CT results or CT findings of unequivocal noninfectious origin. This was modeled on LI-RADS, in which cases that have no nodules or that have nodules with definitely benign features are reported together, as opposed to BI-RADS, in which category 1 refers to normal findings only, we consider this approach more suitable for patients with potential COVID-19; concomitant findings are frequent in the lung, and there is considerable interobserver variability regarding

Table 1: Overview of CO-RADS Categories and the Corresponding Level of Suspicion for Pulmonary Involvement in COVID-19

CO-RADS Category	Level of Suspicion for Pulmonary Involvement of COVID-19	Summary
0	Not interpretable	Scan technically insufficient for assigning a score
1	Very low	Normal or noninfectious
2	Low	Typical for other infection but not COVID-19
3	Equivocal/unsure	Features compatible with COVID-19 but also other diseases
4	High	Suspicious for COVID-19
5	Very high	Typical for COVID-19
6	Proven	RT-PCR positive for SARS-CoV-2

Note.—CO-RADS = COVID-19 Reporting and Data System, COVID-19 = coronavirus disease 2019, RT-PCR = reverse transcription-polymerase chain reaction, SARS-CoV-2 = severe acute respiratory syndrome coronavirus 2.

which findings are normal. According to our definition, mild or severe emphysema, perifissural nodules, lung tumors, and fibrosis are classified as CO-RADS category 1 findings. This category is identical to the "negative for pneumonia" category of the RSNA consensus statement (13).

CO-RADS Category 2

CO-RADS category 2 implies a low level of suspicion for pulmonary involvement by COVID-19 based on CT findings in the lungs typical of infectious origin that are considered not compatible with COVID-19. Examples are bronchitis, infectious bronchiolitis, bronchopneumonia, lobar pneumonia, and pulmonary abscess. Features include tree-in-bud sign, a centrilobular nodular pattern, lobar or segmental consolidation, and lung cavitation. These features are similar to the ones in the "atypical appearance" category of the RSNA consensus statement (13). Cases with smooth interlobular septal thickening with pleural effusion, which is also part of this RSNA category, are assigned to CO-RADS category 1 if considered typical for interstitial pulmonary edema or are assigned to CO-RADS category 3 if ground-glass opacities that may mimic pulmonary involvement by COVID-19 are also present. This choice was made because CO-RADS describes the pulmonary, not cardiac, involvement of COVID-19.

CO-RADS Category 3

CO-RADS category 3 implies equivocal findings for pulmonary involvement of COVID-19 based on CT features that can also be found in other viral pneumonias or noninfectious causes. Findings include perihilar ground-glass opacity, homogenous extensive ground-glass opacity with or without sparing of some secondary pulmonary lobules, or ground-glass opacity together with smooth interlobular septal thickening with or without pleural effusion in the absence of other typical CT findings. CO-RADS category 3 also includes small ground-glass opacities that are not centrilobular (otherwise they would be CO-RADS category 2) or not located close to the visceral pleura (otherwise they would be CO-RADS category 4). In addition, it contains patterns of consolidation compatible with organizing pneumonia without other typical findings of COVID-19. This category partially overlaps with the indeterminate appearance category of the RSNA consensus statement but includes those cases with lower likelihood for COVID-19 (13).

CO-RADS Category 4

CO-RADS category 4 implies a high level of suspicion for pulmonary involvement by COVID-19 based on CT findings that are typical for COVID-19 but also show some overlap with other (viral) pneumonias. Findings are similar to those for CO-RADS category 5; however, they are not in contact with the visceral pleura, nor are they located strictly unilaterally in a predominant peribronchovascular distribution or superimposed on severe diffuse preexisting pulmonary abnormalities. CO-RADS category 4 consists of the features of the indeterminate appearance category of the RSNA consensus statement that are associated with a higher likelihood of COVID-19 (13).

CO-RADS Category 5

CO-RADS category 5 implies a very high level of suspicion for pulmonary involvement by COVID-19 based on typical CT findings (Table 2). Mandatory features are ground-glass opacities with or without consolidations in lung regions close to visceral pleural surfaces, including the fissures, and a multifocal bilateral distribution. Other classifications only describe a peripheral location, but we found that the vicinity to the minor or major fissure is also typical. Subpleural sparing can be present. We found that the previously described lower lobe predominance is frequently not present in otherwise typical RT-PCR–positive cases; therefore, lower lobe predominance was excluded as a required feature.

CO-RADS category 5 requires the presence of at least one confirmatory pattern that aligns with the temporal evolution of the disease (15). The pattern that has been described early in the course of COVID-19 is dominated by multiple ground-glass areas, which often show (half) rounded and unsharp demarcation but can be accompanied by sharply delineated ground-glass areas that outline the shape of multiple adjacent secondary pulmonary lobules. The crazy paving pattern, which has been described as appearing later in the course of the disease, shows visible intralobular lines. As the disease progresses, more consolidations occur within the areas of ground-glass opacity. Finally, opacities that resemble organizing pneumonia occur, such as reverse halo signs or ground-glass opacity with extensive subpleural consolidations and air bronchograms. Subpleural curvilinear bands or bands of ground glass with or without consolidation in a

Table 2: Typical Features for Pulmonary Involvement of COVID-19

Obligatory Features
Ground-glass opacities, with or without consolidations, in lung regions close to visceral pleural surfaces, including the fissures (subpleural sparing is allowed) and multifocal bilateral distribution
Confirmatory Patterns
Ground-glass regions
Unsharp demarcation, (half) rounded shape
Sharp demarcation, outlining the shape of multiple adjacent secondary pulmonary lobules
Crazy paving
Patterns compatible with organizing pneumonia
Thickened vessels within parenchymal abnormalities found in all confirmatory patterns

tethered arching pattern with small connections to the pleura are also considered typical findings. Thickened vessels within lung abnormalities are typical and are frequently found in all other confirmatory patterns. CO-RADS category 5 is largely identical to the typical appearance of the RSNA consensus statement (13).

CO-RADS Category 6

CO-RADS category 6, similar to BI-RADS category 6, was introduced to indicate proven COVID-19, as signified by positive RT-PCR test results for virus-specific nucleic acid.

Materials and Methods

Study Group

An observer study was performed on a set of 105 randomly selected chest CT scans obtained in a group of consecutive patients who presented to the emergency department between March 14, 2020, and March 25, 2020, with suspected SARS-CoV-2 infection in whom RT-PCR was performed. Patient inclusion criteria, CT protocol, and radiation parameters are described in Appendix E2 and Figure E6 (online). Medical ethics committee approval was obtained prior to the study. Informed consent was waived, and data collection and storage were performed in accordance with local guidelines.

Patient characteristics (age, sex, comorbidities); clinical follow-up including a multidisciplinary clinical diagnosis, if applicable; and RT-PCR results were extracted from electronic patient records. These data enabled stratification of all patients into one of the following three groups: patients with at least one positive RT-PCR result for SARS-CoV-2 within 5 days after CT (PCR positive), patients with one or more negative RT-PCR results but a clinical diagnosis of COVID-19 according to clinical records (PCR negative, clinically positive), or patients with one or more negative RT-PCR results and a clinical course not consistent with COVID-19 or consistent with an alternative diagnosis (PCR negative, clinically negative).

CT Scoring Procedure

CT images were extracted from the picture archive and communication system, anonymized, and imported into a browser-based dedicated viewing system for CT scans (CIRRUS Core; *https://grand-challenge.org/reader-studies/*). The software facilitated reading and scoring of anonymized CT images in the three orthogonal views, providing reading tools such as average or maximum intensity projections, window width–window level adaptation, panning, and zooming.

Table 3: Baseline Characteristics of the Patient Study Group

Parameter	Value (n = 105)
Age (y)*	62 ± 16
Male sex	61 (58)
Comorbidities	
Diabetes	15 (14)
Lung disease	41 (39)
Cancer	22 (21)
Immune deficiency	17 (16)
Cardiovascular	46 (44)
Duration of symptoms (d)†	6 (2–10)
No. of RT-PCR assays	
1	84 (80)
2	14 (13)
3	4 (4)
4	2 (2)
5	1 (1)
Positive RT-PCR results	53 (51)

Note.—Unless otherwise indicated, data are numbers of patients and data in parentheses are percentages. RT-PCR = reverse transcription-polymerase chain reaction.

* Data are mean ± standard deviation.

† Data are median, with the interquartile range (in parentheses).

Eight observers from seven hospitals in the Netherlands participated in the study (B.G., J.K., L.F.B., M.P., H.A.G., J.L.S., C.S., T.R.V.). Four observers (B.G., J.K., J.L.S., T.R.V.) had less than 5 years of experience in reading chest CT scans, whereas the others (L.B., M.P., H.G., C.S.) had 5–27 years. All observers were familiar with the CO-RADS score from clinical experience interpreting at least 30 CT scans. Observers assigned CO-RADS scores using scoring software with drop-down lists, with blinding to RT-PCR results and patient information except for age and sex. In addition, they were blinded for the prevalence of COVID-19 in the selected cohort, medical history, and clinical follow-up.

Table 4: CO-RADS Interobserver Comparison and Performance

Observer No.	κ Value	AUC vs RT-PCR	AUC vs Diagnosis
1	0.58 (0.47, 0.69)	0.93 (0.88, 0.98)	0.96 (0.93, 0.99)
2	0.63 (0.52, 0.74)	0.92 (0.86, 0.97)	0.95 (0.91, 0.99)
3	0.68 (0.57, 0.78)	0.90 (0.84, 0.96)	0.94 (0.90, 0.99)
4	0.45 (0.34, 0.57)	0.88 (0.81, 0.95)	0.92 (0.87, 0.98)
5	0.48 (0.37, 0.59)	0.92 (0.87, 0.98)	0.96 (0.93, 1.00)
6	0.65 (0.55, 0.76)	0.93 (0.88, 0.98)	0.96 (0.93, 1.00)
7	0.61 (0.51, 0.72)	0.92 (0.87, 0.98)	0.96 (0.93, 1.00)
8	0.60 (0.49, 0.71)	0.87 (0.80, 0.94)	0.92 (0.86, 0.97)
Overall	0.47 (0.45, 0.49)*	0.91 (0.85, 0.97)	0.95 (0.91, 0.99)

Note.—The κ characteristics of coronavirus disease 2019 Reporting and Data System of each observer are compared with the median of the other observers. Area under the receiver operating characteristic curve (AUC) for each observer is given, separated for the reference standards defined by reverse transcription-polymerase chain reaction (RT-PCR) alone and RT-PCR together with clinical diagnosis. Data in parentheses are 95% confidence intervals.

* Fleiss κ value.

Statistical Analysis

Statistical analysis was performed using SPSS, version 25 (IBM, Armonk, NY). Data are presented as the mean ± standard deviation or median and interquartile range based on normality of data. A 5 × 5 confusion matrix was made separately per observer, in which the CO-RADS score of the observer was compared with the median CO-RADS score of the remaining seven observers. A similar matrix was calculated using the sum of all individual 5 × 5 tables.

For each observer, a receiver operating characteristics curve was calculated, and the area under the receiver operating characteristics curve (AUC) was used to assess the performance of CO-RADS relative to two reference standards for the diagnosis of COVID-19: a positive RT-PCR test (PCR positive) and a reference that combined the results of the RT-PCR test with a clinical COVID-19 diagnosis (PCR positive and PCR negative but clinically positive). Mean AUC across observers and 95% confidence intervals (CIs) were calculated. In addition, the average percentage of cases assigned to each CO-RADS category, including the 95% CI, were determined for *(a)* PCR-positive, *(b)* PCR-negative and clinically positive, and *(c)* PCR-negative and clinically positive cases.

To quantify interobserver agreement, the Fleiss κ value was determined across observers. The κ values were obtained by comparing the CO-RADS scores of each observer to the median score of the remaining seven observers. Interobserver agreement was considered slight for a κ value of 0.01–0.20, fair for a κ value of 0.21–0.40, moderate for a κ value of 0.41–0.60, substantial for a κ value of 0.61–0.80, and almost perfect for a κ value of 0.81–1.00 (16).

Results

Table 3 depicts baseline characteristics of the 105 included patients (mean age, 62 years ± 16; 61 men, 53 with positive RT-PCR results). In 21 patients, at least one repeat RT-PCR assay was performed because of high clinical suspicion for COVID-19 but a negative result at initial RT-PCR. An additional five patients had a clinical diagnosis of COVID-19 despite one (n = 2) to five negative RT-PCR test results (PCR negative, clinically positive).

Interobserver Variability of CO-RADS

There was absolute agreement in assigned CO-RADS category in 573 of 840 (68.2%) observations. A discrepancy by one CO-RADS category was seen in 235 of 840 (28.0%) observations, and of these pairs, CO-RADS categories 4 and 5 and CO-RADS categories 1 and 2 occurred in 128 of 840 (15.2%) observations. A difference of two CO-RADS categories was found in 31 (3.7%) observations, and a difference of three categories was found in one (0.1%) observation. The resulting 5 × 5 table is given in Appendix E3 (online). The Fleiss κ value of all observers on CO-RADS was 0.47 (95% CI: 0.45, 0.49). The κ values and 95% CIs for the individual categories were as follows: CO-RADS 1, 0.58 (95% CI: 0.54, 0.62); CO-RADS 2, 0.36 (95% CI: 0.32, 0.40); CO-RADS 3, 0.31 (95% CI: 0.28, 0.35); CO-RADS 4, 0.20 (95% CI: 0.17, 0.24); CO-RADS 5, 0.68 (95% CI: 0.65, 0.72). The κ value for each observer is provided in Table 4. Agreements of individual observers with the median of the remaining observers were either substantial (n = 4) or moderate (n = 4).

Diagnostic Performance of CO-RADS

CO-RADS was able to distinguish between patients with positive PCR results from those with negative PCR results with an average AUC of 0.91 (95% CI: 0.85, 0.97). Average AUC increased to 0.95 (95% CI: 0.91, 0.99) if a clinical diagnosis of COVID-19 was also accepted (Table 4). The proportion of cases with positive PCR results or clinical diagnosis of COVID-19 increased from CO-RADS 1 to 5, as shown in Figure 1. All five CT scans of patients with negative PCR results but positive clinical findings were assigned CO-RADS categories 3 to 5.

False-negative CO-RADS category 1 relative to a combined clinical and RT-PCR reference standard was found in nine of 161 (5.6%; 95% CI: 1.0%, 10%) ratings. This occurred in four patients: In one patient, CT revealed bronchial wall thickening and very subtle ground-glass opacities in both lower lobes (one observer). In one patient, CT revealed multiple ground-glass opacities in only the right lower lobe and a pleural lesion in the right upper lobe (two observers). In two patients, CT revealed concurrent preexisting disease (hypersensitivity pneumonitis, n = 1; silicosis, n = 1) (six observers). CO-RADS category 2 assessments were false-negative for COVID-19 in 22 of 159 ratings (average, 13.8%; 95% CI: 9%, 18%) in 13 cases, including

all four false-negative cases with a CO-RADS 1 score. Correspondingly, six of 97 (6.2%; 95% CI: 1.5%, 10%) CO-RADS 4 observations and one of 286 (0.3%; 95% CI: 0%, 1.0%) CO-RADS 5 observations were false-positive, as assessed by six observers for four cases. RT-PCR and clinical diagnosis were almost equally distributed among observations of CO-RADS category 3; 57 of 137 (41.6%; 95% CI: 31%, 54%) observations of CO-RADS category 3 were positive for COVID-19. Figure 2 shows examples of RT-PCR–positive and RT-PCR–negative CO-RADS category 3 observations, and examples of CO-RADS category 4 and CO-RADS category 5 observations are shown in Figures 3 and 4.

Discussion

The coronavirus disease 2019 (COVID-19) Reporting and Data System (CO-RADS) was developed as a categorical system to assess suspicion of lung involvement by COVID-19 on chest CT scans and to enable standardized communication. With soaring case numbers and increasing logistic constraints, CO-RADS was readily embraced by clinicians for ease of communication and workflow optimization. CO-RADS is used to assess suspicion for pulmonary involvement at CT. Accordingly, it must be interpreted together with the duration and type of symptoms, as well as clinical and laboratory findings, when it comes to building a clinical diagnosis of COVID-19 with or without lung involvement before reverse transcriptase-polymerase chain reaction tests are available.

Our evaluation of a random sample of patients with symptoms suggestive of COVID-19 showed moderate to substantial interobserver agreement, even though all observers were from different hospitals and had different levels of exposure to CT in patients with COVID-19. Observer agreement was highest in CO-RADS categories 1 and 5, with κ values of 0.58 and 0.68, respectively. In 68.2% of observations, there was absolute agreement of scores, and in 15.2% of observations, scores varied between CO-RADS categories 1 and 2 or between CO-RADS categories 4 and 5, indicating that in more than 80% of the cases, observers agreed on the suspicion for pulmonary involvement of COVID-19 being either low to very low or high to very high. The CO-RADS level of interobserver agreement (Fleiss κ = 0.47) lies between interobserver agreement values for the PI-RADS (Fleiss κ = 0.24) and LI-RADS (Fleiss κ = 0.67) (17,18).

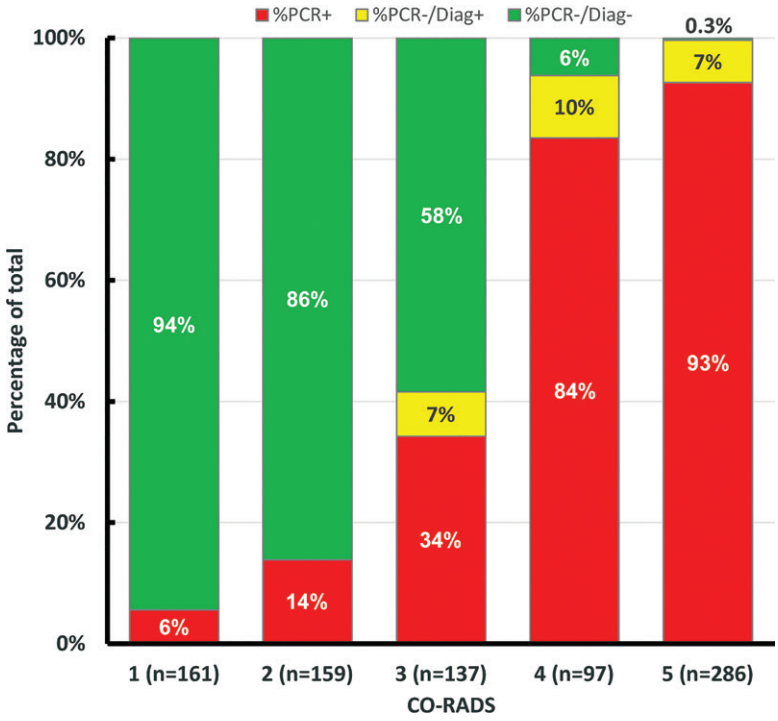

Figure 1: Cumulative coronavirus disease 2019 (COVID-2019) Reporting and Data System (CO-RADS) score versus reverse transcription-polymerase chain reaction (RT-PCR) results and clinical diagnosis. Red columns show cases with a positive RT-PCR result. Yellow columns represent cases with a negative RT-PCR result but a clinical diagnosis of COVID-19. Green columns show the percentage of cases with a negative RT-PCR result for severe acute respiratory syndrome coronavirus 2 and no clinical COVID-19 diagnosis.

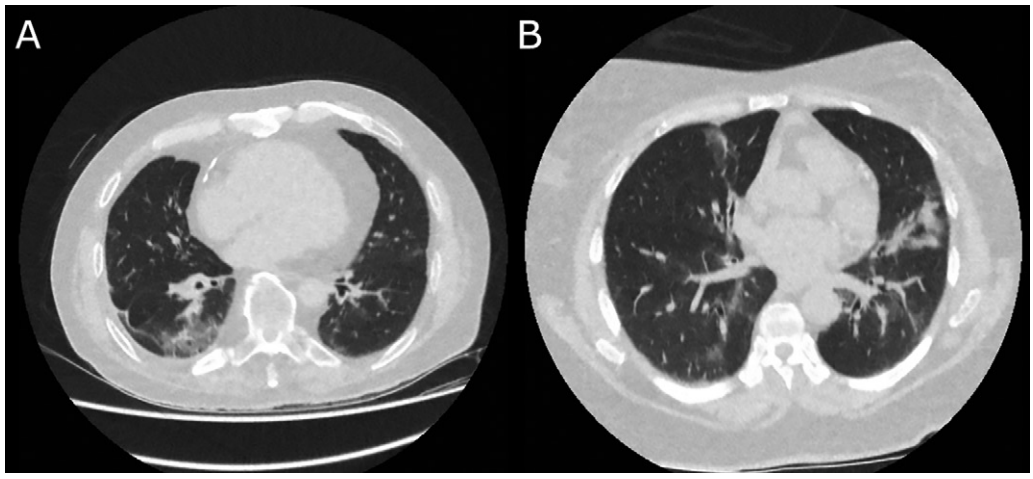

Figure 2: Axial CT slices of the basal lungs in two patients with a majority of coronavirus disease 2019 Reporting and Data System category 3 observations. *A*, A 72-year-old man with a history of cardiovascular disease and chronic obstructive pulmonary disease who presented with fever and a productive cough after 1 day. He had negative reverse transcription-polymerase chain reaction (RT-PCR) test results for severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and a clinical diagnosis of community-acquired pneumonia. He was given antibiotics and discharged after 8 days. *B*, A 63-year-old woman with diabetes, chronic kidney failure, and hypertension who presented with fever and cough after 3 days. RT-PCR results were positive for SARS-CoV-2, and she received oxygen therapy and was discharged after 2 days and advised to quarantine until full resolution of symptoms. Symptoms improved after a few days.

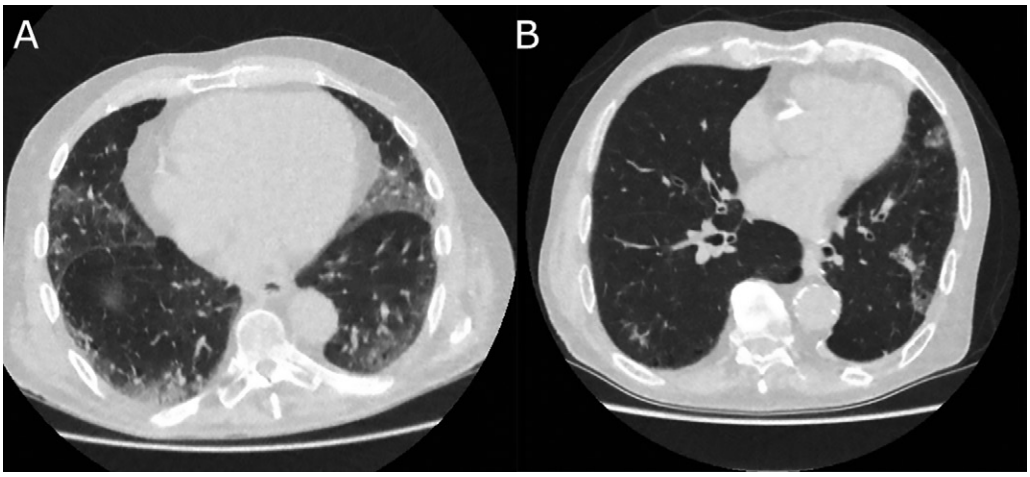

Figure 3: Axial CT slices of the basal lungs in two patients with coronavirus disease 2019 (COVID-19) Reporting and Data System category 4 observations. *A*, A 79-year-old man with a history of pulmonary embolism presented with a cough of 7 days duration, with fever upon fever. The reverse transcription-polymerase chain reaction (RT-PCR) test was positive for severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), for which the patient was admitted and treated. He died despite treatment after 14 days. *B*, A 78-year-old man with a history of chronic obstructive pulmonary disease, lung cancer, and hypertension presented with a productive cough and dyspnea of 5 days duration, with fever upon presentation. The initial RT-PCR test was negative for SARS-CoV-2, but a clinical diagnosis of COVID-19 was made based on typical symptoms, CT characteristics, and absence of an alternative diagnosis. The patient was discharged and advised to stay in quarantine until full resolution of symptoms.

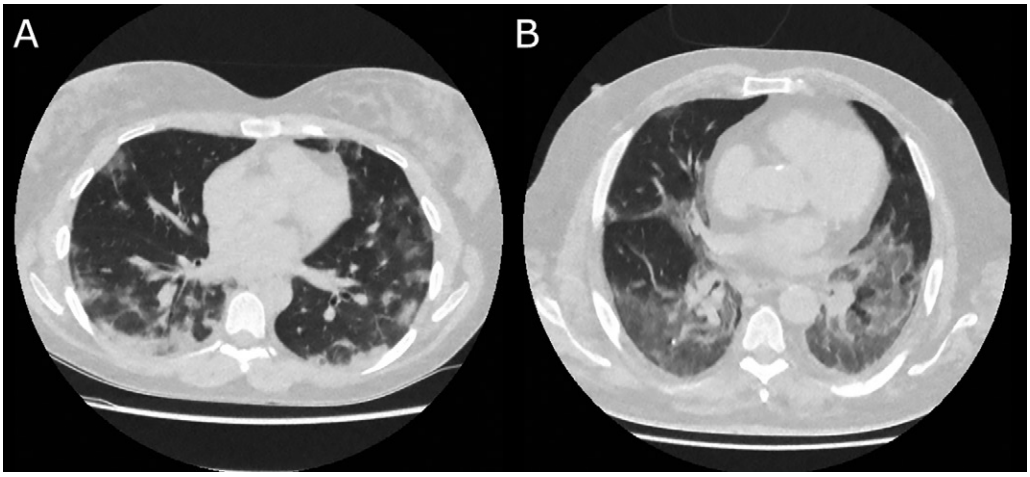

Figure 4: Axial CT slices of the basal lungs in two patients with coronavirus disease 2019 (COVID-19) Reporting and Data System category 5 observations by all eight observers. *A*, A 30-year-old woman with reverse transcription-polymerase chain reaction test results positive for severe acute respiratory syndrome coronavirus 2 who presented with fever and cough of 12 days duration. She was admitted to the COVID-19 ward. She was discharged after 7 days, with resolution of symptoms. *B*, A 51-year-old man presented after 8 days of fever, dyspnea, and cough. A clinical diagnosis of COVID-19 was made due to clinical symptoms and laboratory findings, despite negative results at repeated RT-PCR tests. This patient was admitted for 2 days due to hypoxia, with alleviation of symptoms after 5 days.

In the current setting with a high pretest probability of disease in the acute phase of the pandemic, the performance of CO-RADS was very good, with an average AUC of 0.91 when compared with RT-PCR and an AUC of 0.95 when compared with a combination of RT-PCR and clinical reference standard. However, our results also indicate that the diagnosis of COVID-19 at CT remains difficult in a subset of patients, which emphasizes the importance of a reporting tool that includes diagnostic confidence. Bernheim et al found that CT results can be negative in the early stages of COVID-19 (19), which might have been the case in 13 of 58 patients with COVID-19 whose CT scans were rated as CO-RADS 1 or 2 by at least one observer. Thus, a CO-RADS score of 1 or 2 should be interpreted with caution within the 1st days of disease presence. A CO-RADS score of 3 encompasses a category in which CT alone offers little for the diagnosis of COVID-19. Presumably, knowledge of the prevalence of disease within the patient population, prior imaging studies, or a higher level of experience may decrease the number of equivocal calls. Although CT findings are not specific for COVID-19 (12), they appear highly suggestive of the disease, which is emphasized by only one false-positive rating out of 286 CO-RADS 5 ratings.

There are several caveats concerning the performance of CO-RADS, mainly because the system was developed in the acute stage of the COVID-19 pandemic with rapidly rising case numbers and a parallel restriction in resources. Whether its accuracy remains high in other settings may depend on the prevalence of the disease, the duration of the pandemic, and the prevalence of other diseases with overlapping CT morphology. CO-RADS was developed in a high-prevalence setting, which implies that the positive predictive value is much higher than in a low-prevalence situation. Also, no patients with residual abnormalities, such as subpleural banding after previous COVID-19 infection, existed at the beginning of the pandemic. At the time that SARS-CoV-2 spread throughout Europe, the influenza season was ending, thereby reducing the number of overlapping patterns due to other viruses. Finally, whether this system is sufficient for patients with mild or no symptoms has not been validated.

Our observer study had limitations. First, the study group was relatively small. Second, it was representative of a population presenting to the emergency department in the acute phase

of the SARS-CoV-2 outbreak and requiring hospital admission for clinical reasons. This increases the disease prevalence substantially over that in a population with fewer symptoms. Third, observers had limited experience compared with the experience of physicians in areas with a larger outbreak of SARS-CoV-2. Finally, the diagnosis of COVID-19 was based on clinical decision despite negative RT-PCR results, but this occurred in a small subset of patients (n = 5). In this multidisciplinary decision, the results of the CT scan were known, introducing an affirmation bias. Nevertheless, we included those patients in the study because it reflected current clinical practice.

The coronavirus disease 2019 (COVID-19) Reporting and Data System (CO-RADS), developed by the Dutch Radiological Society, provides a framework that builds on other reporting schemes for COVID-19 but expands the concept in a way similar to systems like LI-RADS. Categories 1–5 provide increasing suspicion for pulmonary involvement of COVID-19 at unenhanced chest CT, thus allowing for task-specific cutoff points for clinical decision making. It provides very good performance in predicting COVID-19 in patients with moderate to severe symptoms and has substantial interobserver agreement, especially for CO-RADS categories 1 and 5. Thus, the system fulfills the need for a structured and fast reporting system that decreases ambiguity in communications with referring physicians and facilitates collection of CT performance data for further research of this worldwide health care problem.

Acknowledgments: Pieternel van der Tol and Willem Jan van der Woude are acknowledged for their help with the chest CT protocol and for collecting radiation dose parameters. Ioannis Sechopoulos, PhD, is acknowledged for proofreading the manuscript. Barbara Janssen , PhD, and Karlijn Groenen , PhD, are acknowledged for their work to obtain ethics board approval.

Author contributions: Guarantors of integrity of entire study, M.P., W.v.E., H.Q.v.U., C.S.P., M.B.; study concepts/study design or data acquisition or data analysis/interpretation, all authors; manuscript drafting or manuscript revision for important intellectual content, all authors; approval of final version of submitted manuscript, all authors; agrees to ensure any questions related to the work are appropriately resolved, all authors; literature research, M.P., W.v.E., T.v.R.V., H.Q.v.U., L.B., H.G., C.S.P., B.v.G., M.B.; clinical studies, M.P., W.v.E., T.v.R.V., H.Q.v.U., L.S., H.G., M.B.; statistical analysis, W.v.E., C.S.P., B.v.G.; and manuscript editing, M.P., W.v.E., T.v.R.V., H.Q.v.U., L.S., L.B., H.G., J.K., C.S.P., B.v.G., M.B.

Disclosures of Conflicts of Interest: M.P. Activities related to the present article: disclosed no relevant relationships. Activities not related to the present article: receives grants and personal fees from Canon Medical Systems and Siemens Healthineers; receives personal fees from Bracco and Bayer. Other relationships: disclosed no relevant relationships. **W.v.E.** disclosed no relevant relationships. **T.v.R.V.** disclosed no relevant relationships. **H.Q.v.U.** disclosed no relevant relationships. **L.S.** disclosed no relevant relationships. **L.B.** disclosed no relevant relationships. **B.G.** disclosed no relevant relationships. **H.G.** disclosed no relevant relationships. **J.K.** disclosed no relevant relationships. **C.S.** Activities related to the present article: disclosed no relevant relationships. Activities not related to the present article: receives royalties from Elsevier, Springer, and Thieme; developed educational presentations for Springer. Other relationships: disclosed no relevant relationships. **B.v.G.** Activities related to the present article: disclosed no relevant relationships. Activities not related to the present article: receives royalties from Delft Imaging, Thirona, and MeVis Medical Solutions; holds stock in Thirona. Other relationships: disclosed no relevant relationships. **M.B.** Activities related to the present article: disclosed no relevant relationships. Activities not related to the present article: received a grant from and gave lectures from Canon Medical Systems. Other relationships: disclosed no relevant relationships.

References

1. Li Y, Yao L, Li J, et al. Stability issues of RT-PCR testing of SARS-CoV-2 for hospitalized patients clinically diagnosed with COVID-19. J Med Virol Published online March 26, 2020. Accessed April 17, 2020.
2. Ai T, Yang Z, Hou H, et al. Correlation of chest CT and RT-PCR testing in coronavirus disease 2019 (COVID-19) in China: a report of 1014 cases. Radiology Published online February 26, 2020. Accessed April 8, 2020.
3. Wang W, Xu Y, Gao R, et al. Detection of SARS-CoV-2 in different types of clinical specimens. JAMA Published online March 11, 2020. Accessed April 17, 2020.
4. Zou L, Ruan F, Huang M, et al. SARS-CoV-2 viral load in upper respiratory specimens of infected patients. N Engl J Med 2020;382(12):1177–1179.
5. Corman VM, Landt O, Kaiser M, et al. Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. Euro Surveill 2020;25(3).
6. Fang Y, Zhang H, Xie J, et al. Sensitivity of chest CT for COVID-19: comparison to RT-PCR. Radiology Published online February 19, 2020. Accessed April 15, 2020.
7. Long C, Xu H, Shen Q, et al. Diagnosis of the coronavirus disease (COVID-19): rRT-PCR or CT? Eur J Radiol 2020;126:108961.
8. Novel Coronavirus Pneumonia Diagnosis and Treatment Plan (Provisional 7th Edition), 04-03-2020. Chinese: http://www.gov.cn/zhengce/zhengceku/2020-03/04/5486705/files/ae61004f930d47598711a0d4cbf874a9.pdf. English translation: https://www.chinalawtranslate.com/en/coronavirus-treatment-plan-7/. Accessed April 16, 2020.
9. Hope MD, Raptis CA, Shah A, Hammer MM, Henry TS; six signatories. A role for CT in COVID-19? What data really tell us so far. Lancet 2020;395(10231):1189–1190.
10. American College of Radiology. ACR recommendations for the use of chest radiography and computed tomography (CT) for suspected COVID-19 infection. March 11, 2020. https://www.acr.org/Advocacy-and-Economics/ACR-Position-Statements/Recommendations-for-Chest-Radiography-and-CT-for-Suspected-COVID19-Infection. Updated March 22, 2020. Accessed April 17, 2020.
11. Rubin GD, Ryerson CJ, Haramati LB, et al. The role of chest imaging in patient management during the COVID-19 pandemic: a multinational consensus statement from the Fleischner Society. Radiology Published online April 7, 2020. Accessed April 15, 2020.
12. Bai HX, Hsieh B, Xiong Z, et al. Performance of radiologists in differentiating COVID-19 from viral pneumonia on chest CT. Radiology Published online March 10, 2020. Accessed April 17, 2020.
13. Simpson S, Kay FU, Abbara S, 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. J Thorac Imaging 2020;2(2):e200152.
14. An JY, Unsdorfer KML, Weinreb JC. BI-RADS, C-RADS, CAD-RADS, LI-RADS, Lung-RADS, NI-RADS, O-RADS, PI-RADS, TI-RADS: Reporting and Data Systems. RadioGraphics 2019;39(5):1435–1436.
15. Wang Y, Dong C, Hu Y, et al. Temporal Changes of CT Findings in 90 patients with COVID-19 pneumonia: a longitudinal study. Radiology Published online March 19, 2020. Accessed April 17, 2020.
16. Fleiss JL, Cohen J. The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. Educ Psychol Meas 1973;33(3):613–619.
17. Smith CP, Harmon SA, Barrett T, et al. Intra- and interreader reproducibility of PI-RADSv2: A multireader study. J Magn Reson Imaging 2019;49(6):1694–1703.
18. van Riel SJ, Jacobs C, Scholten ET, et al. Observer variability for Lung-RADS categorisation of lung cancer screening CTs: impact on patient management. Eur Radiol 2019;29(2):924–931.
19. Bernheim A, Mei X, Huang M, et al. Chest CT findings in coronavirus disease-19 (COVID-19): relationship to duration of infection. Radiology Published February 20, 2020. Accessed April 15, 2020.