Diagnostic performance of co-rads and the rsna classification system in evaluating covid-19 at chest cta meta-analysis

Kwee, Robert M.; Adams, Hugo J.A.; Kwee, Thomas C.

Published in:
Radiology: cardiothoracic imaging

DOI:
10.1148/ryct.2021200510

Radiology: Cardiothoracic Imaging

ORIGINAL RESEARCH

Diagnostic Performance of CO-RADS and the RSNA Classification System in Evaluating COVID-19 at Chest CT: A Meta-Analysis

Robert M. Kwee, MD, PhD • *Hugo J. A. Adams, MD, PhD* • *Thomas C. Kwee, MD, PhD*

From the Department of Radiology, Zuyderland Medical Center, Henri Dunantstraat 5, 6419 PC Heerlen, Heerlen/Sittard/Geleen, the Netherlands (R.M.K.); and Department of Radiology, Nuclear Medicine and Molecular Imaging University Medical Center Groningen, University of Groningen, Groningen, the Netherlands (T.C.K.). Received September 2, 2020; revision requested October 10; revision received December 3; accepted January 4, 2021. **Address correspondence to** R.M.K. (e-mail: *rmkwee@gmail.com*).

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

Radiology: Cardiothoracic Imaging 2021; 3(1):e200510 • https://doi.org/10.1148/ryct.2021200510 • Content code: CH

Purpose: To determine the diagnostic performance of the COVID-19 Reporting and Data System (CO-RADS) and the Radiological Society of North America (RSNA) categorizations in patients with clinically suspected coronavirus disease 2019 (COVID-19) infection.

Materials and Methods: In this meta-analysis, studies from 2020, up to August 24, 2020, were assessed for inclusion criteria of studies that used CO-RADS or the RSNA categories for scoring chest CT in patients suspected of having COVID-19. A total of 186 studies were identified. After review of abstracts and text, a total of nine studies were included in this study. Patient information (*n*, age, sex), CO-RADS and RSNA scoring categories, and other study characteristics were extracted. Study quality was assessed with the Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2) tool. Meta-analysis was performed with a random effects model.

Results: Nine studies (3283 patients) were included. Overall study quality was good, except for risk of nonperformance of repeated reverse-transcriptase polymerase chain reaction (RT-PCR) testing after negative initial RT-PCR testing and persistent clinical suspicion in four studies. Pooled COVID-19 frequencies in CO-RADS categories were: 1, 8.8%; 2, 11.1%; 3, 24.6%; 4, 61.9%; and 5, 89.6%. Pooled COVID-19 frequencies in RSNA classification categories were: negative, 14.4%; atypical, 5.7%; indeterminate, 44.9%; and typical, 92.5%. Pooled pairs of sensitivity and specificity using CO-RADS thresholds were the following: at least 3, 92.5% (95% CI: 87.1, 95.7) and 69.2% (95%: CI: 60.8, 76.4); at least 4, 85.8% (95% CI: 78.7, 90.9) and 84.6% (95% CI: 79.5, 88.5); and 5, 70.4% (95% CI: 60.2, 78.9) and 93.1% (95% CI: 87.7, 96.2). Pooled pairs of sensitivity and specificity using RSNA classification thresholds for indeterminate were 90.2% (95% CI: 87.5, 92.3) and 75.1% (95% CI: 68.9, 80.4) and for typical were 65.2% (95% CI: 37.0, 85.7) and 94.9% (95% CI: 86.4, 98.2).

Conclusion: COVID-19 infection frequency was higher in patients categorized with higher CO-RADS and RSNA classification categories.

Supplemental material is available for this article.

The coronavirus disease 2019 (COVID-19) pandemic has caused a major global crisis. On December 2, 2020, there were 64 million confirmed cases and almost 1.5 million confirmed deaths due to COVID-19 worldwide (1). Although most countries have already experienced the first surge of rising COVID-19 cases, second surges have started in late 2020. Chest imaging has an important role in the evaluation of patients with COVID-19 (2). The chest imaging findings of COVID-19 were first reported in January 2020 and included bilateral lung involvement and ground-glass opacities in the majority of hospitalized patients (3). Since this first report (3), several studies on the diagnostic value of chest CT in COVID-19 have been published. However, as most initial studies did not use uniform diagnostic criteria (4), their results cannot directly be translated to clinical practice.

Two major chest CT classification scales for standardized CT reporting of COVID-19 have been developed, namely the COVID-19 Reporting and Data System (CO-RADS) (5) and the Radiological Society of North America (RSNA) classification system for reporting COVID-19 pneumonia (6). CO-RADS basically consists of five categories (CO-RADS 1 to 5; Table E1 and Figs E1–E5 [supplement]), whereas the RSNA classification system consists of four categories (negative, atypical, indeterminate, and typical; Table E2 and Figs E1–E5 [supplement]). CO-RADS and the RSNA chest CT classification system are very similar. CO-RADS categories 1, 2, 3–4, and 5 are essentially equal to categories negative, atypical, indeterminate, and typical of the RSNA classification system, respectively (5,7). The use of these standardized diagnostic classification systems may reduce observer variation, enhance clinical communication, and improve generalizability. However, the diagnostic yields of both the CO-RADS and RSNA categorizations are not completely clear yet. Original studies on this topic may suffer from small sample sizes and potential methodologic quality concerns. Aggregated data are necessary to understand the clinical interpretability of these

Abbreviations

CO-RADS = COVID-19 reporting and data system, COVID-19 = coronavirus disease 2019, QUADAS-2 = Quality Assessment of Diagnostic Accuracy Studies 2, RSNA = Radiological Society of North America, RT-PCR = reverse-transcription polymerase chain reaction

Summary

The frequency of coronavirus disease 2019 infection was higher in patients with higher CO-RADS and RSNA classification categories, which supports the order of grading used by both systems.

Key Points

- Using the lowest clinically meaningful thresholds of CO-RADS of at least 3 and indeterminate according to the RSNA classification, sensitivity values were 92.5% and 90.2%, which implies that CO-RADS 1 and 2 and RSNA classification categories negative and atypical certainly do not exclude COVID-19.
- Using the highest thresholds of CO-RADS 5 and typical according to the RSNA classification, specificity values increased up to 93.1% and 94.9% at the cost of sensitivity, with values of 70.4% and 65.2%, respectively.

chest CT classification systems for the diagnosis of COVID-19. Although there have already been meta-analyses published on the diagnostic performance of chest CT in detecting COVID-19 (4,8), the initial studies included within these meta-analyses suffered from methodologic quality issues and did not use uniform diagnostic criteria such as the CO-RADS and RSNA categorizations. These shortcomings limit translation of diagnostic performance values to clinical practice. Therefore, our objective was to determine, in a meta-analysis, the diagnostic performance of the CO-RADS and the RSNA classification systems in patients with clinically suspected COVID-19 infection.

Materials and Methods

The study adhered to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guideline (9).

Data Sources

A search in MEDLINE and Embase was conducted to find original publications on the diagnostic performance of the CO-RADS and the RSNA classification systems in evaluating symptomatic patients with clinically suspected COVID-19 infection. The following search term was used: (CO-RADS OR CORADS OR Radiological Society of North America OR RSNA) AND (Corona OR Coronavirus OR Covid-19 OR SARS-Cov-2 OR 2019nCoV OR Wuhan-virus) AND (Computed tomography OR Computerized tomography OR Computed tomographic OR CT OR CT OR HRCT).

In addition, the journal *Radiology: Cardiothoracic Imaging* was manually searched for potentially relevant publications. Publications that cited the original CO-RADS (5) and RSNA classification system for reporting COVID-19 pneumonia (6) were also searched using the cited reference function in Web of Science and MEDLINE.

The search was updated until August 24, 2020.

Study Selection

Original studies that provided data on the diagnostic performance of the CO-RADS or RSNA classification system in evaluating patients with clinically suspected COVID-19 infection, and in which reverse-transcription polymerase chain reaction (RT-PCR) was the reference standard, were eligible for inclusion. Reviews, abstracts, and studies were excluded for the following reasons: *(a)* included fewer than 10 patients, *(b)* reported insufficient data to compose a 2×2 contingency table to calculate sensitivity and specificity on per-patient level for any CO-RADS or RSNA classification system threshold, and *(c)* only provided data on the performance of artificial intelligence–based analyses. When overlapping data were presented in more than one study, the study with the largest number of patients was selected. Titles and abstracts of retrieved studies were reviewed using the aforementioned selection criteria. The full-text version of each potentially eligible study was then reviewed to definitively determine whether the study fulfilled the selection criteria.

Study Data Extraction

For each included study, the main characteristics (country of origin, patient inclusion period, number of patients, age, and sex of patients, clinical characteristics of included patients, CT protocol, CT interpreters, reference standard, and COVID-19 frequency) were extracted by two independent reviewers (R.M.K., radiologist, and H.J.A.A., 3rd-year resident in radiology). If data from multiple readers were reported, only data from the first reader were extracted and used for the analyses. The number of patients with and without COVID-19 according to the different CO-RADS and the RSNA classification categories was also extracted. Data on interobserver or intraobserver agreement using the CO-RADS and the RSNA classification system were also extracted. Any discrepancies were solved by consensus with a third reviewer (T.C.K., radiologist).

Study Quality Assessment

The quality of included studies was assessed by two independent reviewers (R.M.K. and H.J.A.A.) using the Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2) tool, which comprises four key items: patient selection, index test, reference standard, and flow and timing (10). Any discrepancies were solved by consensus with a third reviewer (T.C.K.).

Statistical Analyses

Frequency of COVID-19 in each of the categories of the CO-RADS and the RSNA classification system were calculated for each individual study and pooled with a random-effects model. Sensitivity and specificity of the CO-RADS and RSNA classification systems at specific diagnostic thresholds in detecting COVID-19 (ie, CO-RADS thresholds of at least 3, at least 4, 5, and RSNA classification thresholds indeterminate and typical) were pooled using a bivariate random-effects model (11). The numbers were pooled in each CO-RADS and in each RSNA classification category separately. The same random-effects model was used per each study, across different categories.

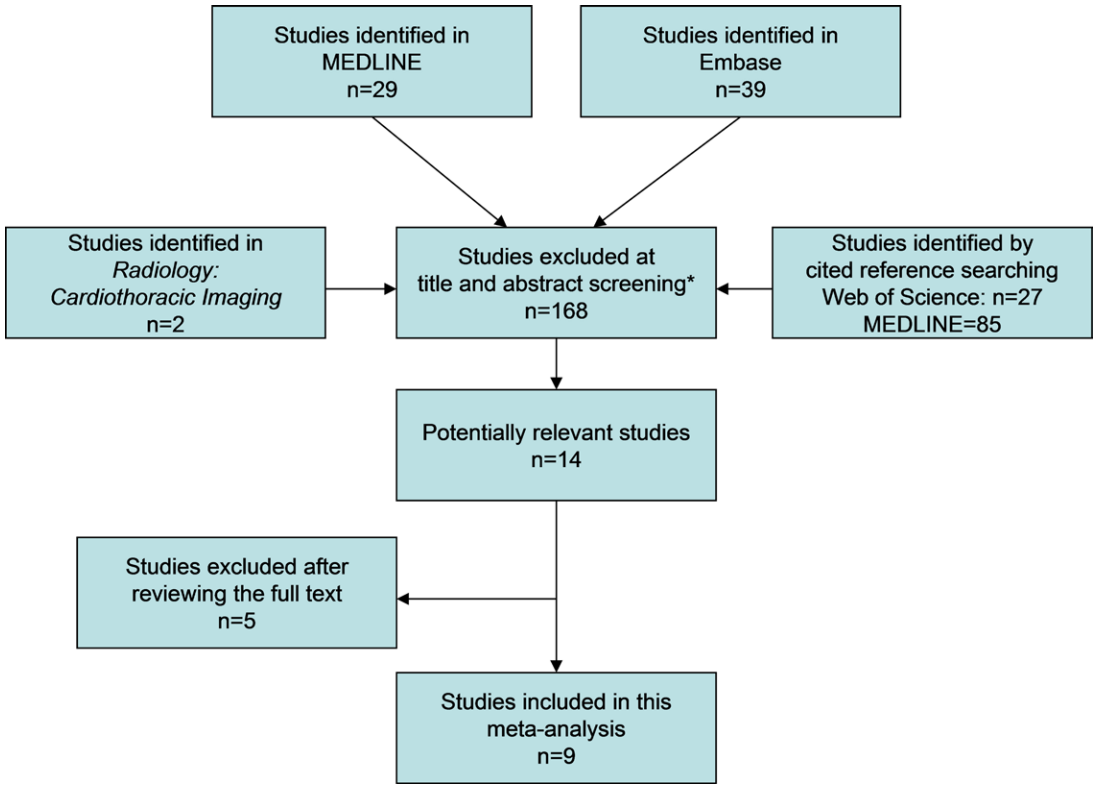

Figure 1: Flow diagram of study selection. The asterisk indicates that there were duplicate studies.

Cochran Q and χ^2 tests were performed to test for heterogeneity between studies, which was defined as $P < .10$. Statistical analyses were performed using the Open Meta-Analyst software package (12) and Meta-analysis of Diagnostic Accuracy Studies package in R software (13,14).

Results

Literature Search

Figure 1 displays the study selection process. A total of 182 studies were eligible for inclusion after searching databases. After screening titles and abstracts, 168 studies were excluded, leaving 14 studies that were potentially eligible for inclusion. After reading the full text of the 14 studies, three studies (15–17) were excluded because the diagnostic performance of either CO-RADS or the RSNA classification system was not investigated, one study (5) was excluded because no data on a per-patient level were reported, and another study (18) was excluded because there were overlapping data with another study (7) which comprised a larger number of patients. Nine studies were eventually included (7,19–26).

The main study characteristics are shown in Table 1 and Table E3 (supplement). All assessed studies were performed between January and June 2020. The median number of patients per study was 312 (range, 71–859), and the total number of patients of all studies combined was 3283. All nine studies included patients with a clinical suspicion of COVID-19. The mean frequency of COVID-19 was 48.7% (range, 41.7%–59.8%). Of all patients included in the nine studies, 1979 patients were evaluated with CO-RADS and 1400 patients were evaluated with the RSNA classification system.

Study Quality

Figure 2 provides a summary of the QUADAS-2 quality assessments. In one study (19), it was unclear whether patients were enrolled consecutively or randomly. There was no risk of bias with regard to patient selection in the other studies or with regard to index test. Risk of bias with respect to reference test was rated high in three studies (22,25,26) because repeated RT-PCR testing was not used in all patients with a negative initial RT-PCR result and persistent clinical suspicion of COVID-19. Risk of bias with respect to reference test was rated unclear in one study (20) because it was not clear whether all patients with an initial negative RT-PCR result and a persistent clinical suspicion of COVID-19 underwent repeated RT-PCR testing. In one study (19), there was potential risk of bias with regard to flow and timing because the time interval between CT and RT-PCR testing was not reported. There was no risk of bias with regard to flow and timing in the other studies because the maximum time interval between chest CT and RT-PCR did not exceed 7 days (21). There were no applicability concerns.

Diagnostic Performance of CO-RADS

The frequency of COVID-19 in each of the categories of CO-RADS is displayed in Table 2. With higher CO-RADS classification, the frequency of COVID-19 increased.

Table 1: Main Characteristics of the Included Studies

Study	Country	Inclusion Period	No. of Patients (Men)	Age (y)	Inclusion Clinical Characteristics	COVID-19 Frequency
CO-RADS						
Fujioka et al (19)	Japan	Jan–Jun	154 (101)	61.3 (21–93)	Symptomatic patients who were suspected by a clinician of having COVID-19 based on symptoms and history of exposure.	49.4% (76/154)
De Smet et al (20)	Belgium	Mar 19– Apr 20	859 (443)	by sex*	WHO-listed symptoms of COVID-19 pneumonia	41.7% (358/859)
Hermans et al (23)	Netherlands	Mar 27–Apr 20	319 (157)	range 44–75	Suspected infection with COVID-19 in combination with at least one of the following†	41.7% (133/319)
Korevaar et al (24)	Netherlands	Mar 16–Apr 16	239 (139)	median, 63 (IQR 51–71)	Suspected COVID-19‡	52.7% (126/239)
RSNA						
Falaschi et al (21)	Italy	Mar 3–Apr 9	773 (424)	62.4 (16–100)	Suspected for COVID-19§	59.8% (462/773)
Ciccarese et al (22)	Italy	Feb 27–Mar 27	460 (267)	54 (14–97)	Suspected with COVID-19 pneumonia‖	45.9% (211/460)
Magalhães Santos et al (25)	Brazil	Mar 13–Mar 23	71 (33)	47.2 (8–94)	Patients who fulfilled the clinical criteria for confirmed COVID-19	50.7% (36/71)
Dofferhoff et al (26)	Netherlands	Mar 8–Mar 31	312 (168)	64 (18–94)	Patients with fever of unknown origin and patients with recent respiratory symptoms with or without fever	49.4% (154/312)
CO-RADS and RSNA						
de Jaegere et al (7)	Netherlands	Mar 12–Mar 23	96 (61)	median 70 (range 29–94)	Clinical suspicion of COVID-19 (ie, fever, cough, and/or shortness of breath)	46.9% (45/96)

Note.—The year for all inclusion dates are in 2020; months and days of the month are shown. Age shown as mean (range) unless otherwise specified. Time interval indicates the time interval between symptom onset and chest CT. CT protocol, time between symptom onset and CT, and information about image interpreters are shown in Table E3 (supplement). COVID-19 = coronavirus disease 2019, IQR = interquartile range.

* Male patients had median age of 71 years (interquartile range, 54–80) and female patients had median age of 68 years (interquartile range, 51–82).

† Criteria were new respiratory symptoms persisting for less than 2 weeks and present during the last 24 hours; saturation of less than 94% and/or respiration rate of greater than 20 breaths per minute and/or abdominal complaints; and a high clinical suspicion even in the absence of symptoms.

‡ Criteria were those with fever, cough or dyspnea, or other signs suggestive of COVID-19 (eg, gastrointestinal symptoms).

§ Criteria were when one or more of these conditions were met: presence of fever (ie, temperature > 37.5°C), cough and dyspnea; presence of mild symptoms and ascertained close contact with a patient with confirmed COVID-19; one previously positive laboratory test result.

‖ Criteria were patients presenting with fever (of unknown origin) or respiratory symptoms.

Pooled frequency of COVID-19 in CO-RADS categories 1, 2, 3, 4, and 5 were 8.8%, 11.1%, 24.6%, 61.9%, and 89.6%, respectively. Pooled sensitivity and specificity of the CO-RADS and the RSNA classification system at specific thresholds are displayed in Table 3. Pooled pairs of sensitivity and specificity using CO-RADS thresholds were the following: at least 3, 92.5% (95% CI: 87.1, 95.7) and 69.2% (95% CI: 60.8, 76.4); at least 4, 85.8% (95% CI: 78.7, 90.9) and 84.6% (95% CI: 79.5, 88.5); and 5, 70.4% (95% CI: 60.2, 78.9) and 93.1% (95% CI: 87.7, 96.2).

Diagnostic Performance of the RSNA Classification System

The frequency of COVID-19 in each of the categories of the RSNA classification systems is displayed in Table 4. With higher RSNA classification, the frequency of COVID-19 increased. Pooled frequencies of COVID-19 in RSNA classification categories negative, atypical, indeterminate, and typical were 14.4%, 5.7%, 44.9%, and 92.5%, respectively. Pooled sensitivity and specificity of the RSNA classification system at specific thresholds are displayed in Table 5. Pooled pairs of sensitivity and specificity using RSNA classification thresholds

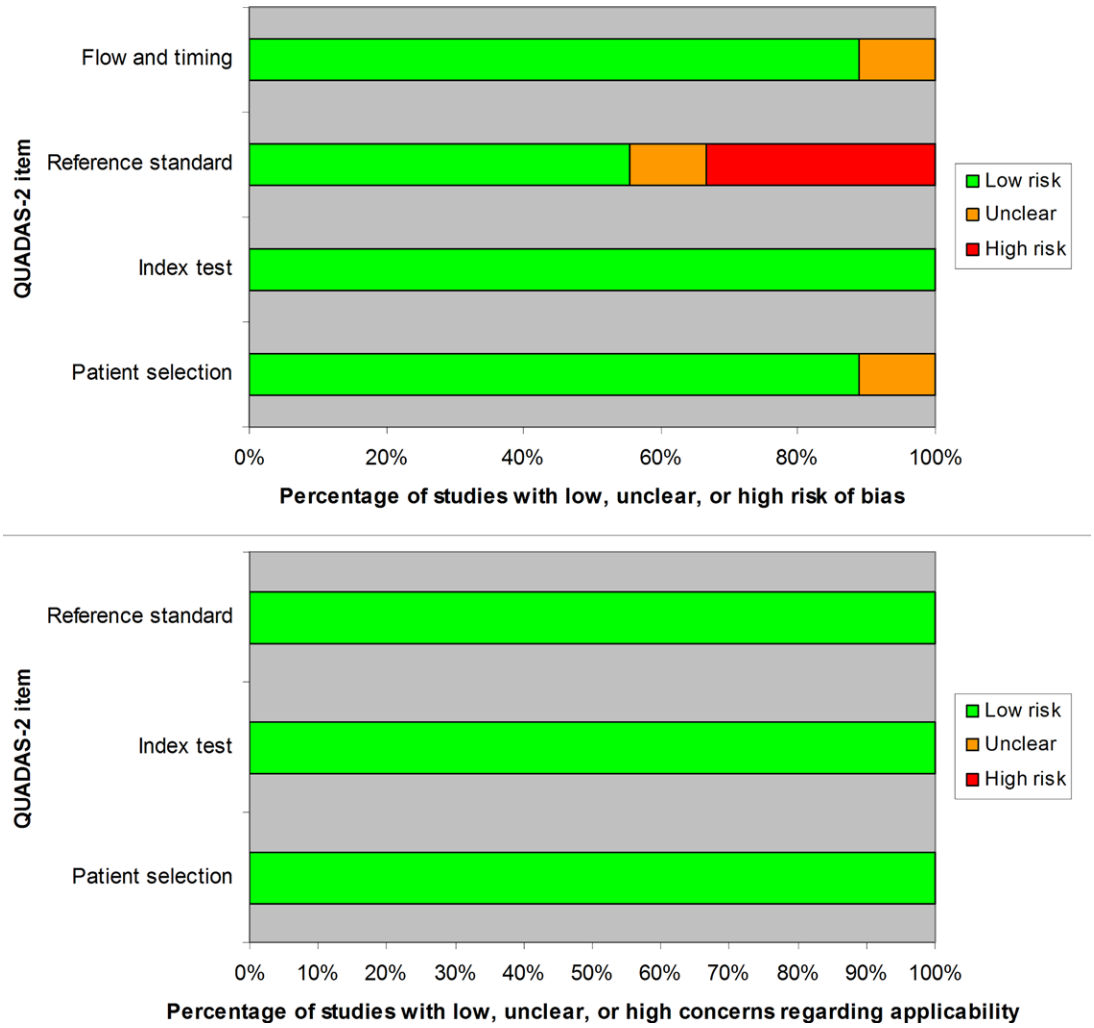

Figure 2: Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2) quality assessments of included studies.

Table 2: Frequency of COVID-19 in each of the Categories of CO-RADS

Study	CO-RADS 1	CO-RADS 2	CO-RADS 3	CO-RADS 4	CO-RADS 5
Fujioka et al (19)*	18.0% (9/50)	28.6% (6/21)	69.2% (9/13)	75.0% (12/16)	90.9% (40/44)
De Smet et al (20)	8.6% (27/313)	13.5% (12/89)	19.5% (15/77)	36.8% (25/68)	89.4% (279/312)
Hermans et al (23)	6.1% (6/99)	9.4% (3/32)	9.1% (4/44)	64.5% (20/31)	90.1% (100/111)
Korevaar et al (24)†	5.9% (4/68)		17.2% (5/29)	82.4% (117/142)	
de Jaegere et al (7)*	11.1% (1/9)	3.1% (1/32)	38.5% (5/13)	76.9% (10/13)	96.6% (28/29)
Dofferhoff et al (26)	10.2% (9/88)	14.3% (3/21)	19.4% (6/31)	63.0% (17/27)	82.1% (119/145)
Pooled frequency‡	8.8% (6.2, 11.4)	11.1% (4.3, 18.0)	24.6% (12.8, 36.5)	61.9% (45.0–78.7)	89.6% (85.6, 93.7)
P value for heterogeneity§	.35	.048	<.001	<.001	.04

Note.—The 95% CI is shown within parenthesis for the pooled frequency.
* Data from the first reader.
† CO-RADS categories 1 and 2, and CO-RADS categories 4 and 5 were merged.
‡ CO-RADS 1, 2, 4, and 5 data from the study of Korevaar et al (24) were not included in the pooled analysis.
§ Statistical heterogeneity between studies was defined as $P < .10$.

Table 3: Pooled Sensitivity and Specificity at Specific Thresholds according to CO-RADS

	Threshold CO-RADS ≥ 3		Threshold CO-RADS ≥ 4		Threshold CO-RADS 5	
Study	Sensitivity	Specificity	Sensitivity	Specificity	Sensitivity	Specificity
Fujioka et al (19)	68.4% (57.3, 77.8)	88.2% (78.5, 93.9)	68.4% (57.3, 77.8)	88.2% (78.5, 93.9)	52.6% (41.6, 63.5)	94.1% (85.8, 97.7)
De Smet et al (20)	84.9% (80.8, 88.3)	84.8% (81.4, 87.7)	84.9% (80.8, 88.3)	84.8% (81.4, 87.7)	77.9% (73.4, 81.9)	93.4% (90.9, 95.3)
Hermans et al (23)	90.2% (84.0, 94.2)	88.2% (82.7, 92.1)	90.2% (84.0, 94.2)	88.2% (82.7, 92.1)	75.2% (67.2, 81.8)	94.1% (89.7, 96.7)
Korevaar et al (24)	92.9% (87.0, 96.2)	77.9% (69.4, 84.5)	92.9% (87.0, 96.2)	77.9% (69.4, 84.5)	No data available	No data available
de Jaegere et al (7)	84.4% (71.2, 92.3)	92.2% (81.5, 96.9)	84.4% (71.2, 92.3)	92.2% (81.5, 96.9)	62.2% (47.6, 74.9)	98.0% (89.7, 99.7)
Dofferhoff et al (26)	88.3% (82.3, 92.5)	77.2% (70.1, 83.1)	88.3% (82.3, 92.5)	77.2% (70.1, 83.1)	77.3% (70.0, 83.2)	83.5% (77.0, 88.5)
Pooled values*	92.5% (87.1, 95.7)	69.2% (60.8, 76.4)	85.8% (78.7, 90.9)	84.6% (79.5, 88.5)	70.4% (60.2, 78.9)	93.1% (87.7, 96.2)
P value for heterogeneity†	<.001	<.001	<.001	.01	<.001	<.001

Note.—Values in parenthesis are the 95% CIs.
* For the pooled analysis, data from the first readers from the study of Fujioka et al (19) and de Jaegere et al (7) were used.
† Statistical heterogeneity between studies was defined as $P < .10$.

Table 4: Frequency of COVID-19 in Each of the Categories of the RSNA Classification System

Study	Negative	Atypical	Indeterminate	Typical
Falaschi et al (21)	14.9% (43/288)*		86.3% (419/485)*	
Ciccarese et al (22)	13.8% (17/123)	10.4% (7/67)	36.7% (36/98)	87.8% (151/172)
Magalhães Santos et al (25)	15.0% (3/20)	0.0% (0/14)	30.0% (3/10)	96.8% (30/31)
de Jaegere et al (7)†	25.0% (2/8)	3.2% (1/31)	64% (25/39)	96.4% (17/18)
Pooled frequency‡	14.4% (8.8, 19.9)	5.7% (0.9, 10.4)	44.9% (24.1, 65.7)	92.5% (86.1, 98.9)
P value for heterogeneity§	.77	.29	.007	.07

* RSNA classification categories negative and atypical, and RSNA classification categories indeterminate and typical were merged.
† Data from the first reader.
‡ Data from the study Falaschi et al (21) were not included in the pooled analysis.
§ Statistical heterogeneity between studies was defined as $P < .10$.

were the following: indeterminate, 90.2% (95% CI: 87.5, 92.3) and 75.1% (95% CI: 68.9, 80.4) and typical, 65.2% (95% CI: 37.0, 85.7) and 94.9% (95% CI: 86.4, 98.2).

Interobserver and Intraobserver Agreement

For the CO-RADS, substantial to almost perfect interobserver agreement has been reported, with κ values of 0.648 to 0.773 (7) and intraclass correlation coefficients of 0.800 to 0.874 (19). For the RSNA classification system, moderate to substantial interobserver agreement has been reported, with κ values of 0.500 (22) and of 0.570 to 0.663 (7). None of the included studies reported data on intraobserver agreement.

Discussion

This meta-analysis provides pooled data with regard to the frequency of patients with COVID-19 for each category of CO-RADS and the RSNA classification system in patients clinically suspected of having COVID-19 infection. With the higher CO-RADS and RSNA classification category, the frequency of patients with COVID-19 increased. This supports the order of grading that is used by both systems. In CO-RADS 5, the prevalence of COVID-19 was 89.6%. In the RSNA category typical, the frequency of COVID-19 was 92.5%. We also provided sensitivity and specificity values for specific diagnostic thresholds. Using the lowest clinically meaningful thresholds of CO-RADS of at least 3 and indeterminate according to the

Table 5: Pooled Sensitivity and Specificity at Specific Thresholds according to the RSNA Classification System

	Threshold Indeterminate		Threshold Typical	
Study	Sensitivity	Specificity	Sensitivity	Specificity
Falaschi et al (21)	90.7% (87.7, 93.0)	78.8% (73.9, 83.0)	No data available	No data available
Ciccarese et al (22)	88.6% (83.6, 92.2)	69.3% (63.5, 74.5)	71.6% (65.1, 77.2)	71.6% (65.1, 77.2)
Magalhães Santos et al (25)	91.7% (78.2, 97.1)	80.5% (66.0, 89.8)	83.3% (68.1, 92.1)	97.4% (86.8–99.5)
de Jaegere et al (7)	93.3% (82.1, 97.7)	73.7% (61.0, 83.4)	37.8% (25.1, 52.4)	98.0% (89.7, 99.7)
Pooled values*	90.2% (87.5, 92.3)	75.1% (68.9, 80.4)	65.2% (37.0, 85.7)	94.9% (86.4, 98.2)
P value for heterogeneity†	.73	.05	<.001	.13

Note.—Values in parenthesis are the 95% CIs.
* For the pooled analysis, data from reader 1 from the study de Jaegere et al (7) were used.
† Statistical heterogeneity between studies was defined as $P < .10$.

RSNA classification, sensitivity values were 92.5% (95% CI: 87.1, 95.7) and 90.2% (95% CI: 87.5, 92.3), respectively. These findings imply that CO-RADS 1 and 2 and RSNA classification categories negative and atypical do not exclude COVID-19. Furthermore, when using these low diagnostic thresholds, specificity is only moderate with values of 69.2% (95% CI: 60.8, 76.4) for CO-RADS of at least 3 and 75.1% (95% CI: 68.9, 80.4) for RSNA indeterminate. If higher diagnostic thresholds are applied, specificity naturally increases at the cost of sensitivity. Using CO-RADS of at least 5 and the RSNA classification typical as diagnostic thresholds, specificity values increased up to 93.1% (95% CI: 87.7, 96.2) and 94.9% (95% CI: 86.4, 98.2). However, when using these high diagnostic thresholds, sensitivity is only moderate with values of 70.4% (95% CI: 60.2, 78.9) and 65.2% (95% CI: 37.0, 85.7).

Methodologic quality of the studies included in the current meta-analysis generally appeared to have higher quality than studies included within prior meta-analyses (4,8). In two prior meta-analyses, high risk of bias was present in all six included studies (100%) (4) and in 10 of 13 included studies (77%) (8). In our current meta-analysis, the reference standard was the only QUADAS-2 item which was deemed to be of high risk of bias. This item applied to three of the nine included studies (33%) because repeated RT-PCR testing was not used in all patients with a negative initial RT-PCR result and persistent clinical suspicion of COVID-19 (22,25,26).

Importantly, we provided a meta-analysis that specifically focused on the diagnostic performance of chest CT in COVID-19 by selecting studies that used standardized diagnostic criteria. Therefore, our study results were more generalizable and useful to clinical practice compared with other prior meta-analyses on CT for COVID-19 assessment. Our finding that CO-RADS 1 and 2 and RSNA classification categories negative and atypical do not exclude COVID-19 was in line with the results of a meta-analysis in nearly 3500 patients, which reported an estimated frequency of 10.6% for normal chest CT findings in symptomatic patients with COVID-19 (27). In a prior meta-analysis of six studies which did not use uniform diagnostic criteria, pooled sensitivity and specificity were 94.6% (95% CI: 91.9, 96.4) and 46.0% (95% CI: 31.9, 60.7), respectively (4). Using CO-RADS of at least 3 and RSNA classification indeterminate as diagnostic thresholds, similar sensitivity values of 92.5% (95% CI: 87.1, 95.7) and 90.2% (95% CI: 87.5, 92.3) can be achieved, while relatively higher specificity values of 69.2% (95% CI: 60.8, 76.4) and 75.1% (95% CI: 68.9, 80.4) are obtained. Thus, when using CO-RADS or the RSNA classification system instead of nonstandardized criteria, it appeared that specificity may be improved without sacrificing sensitivity.

If a low threshold is being used (eg, any lung abnormality on chest CT is considered positive for COVID-19), virtually all COVID-19 cases with lung abnormalities will be correctly classified, but all non–COVID-19 cases with any lung abnormality at chest CT will be incorrectly classified as having COVID-19 (28). By applying standardized diagnostic criteria such as CO-RADS or the RSNA classification system, a higher proportion of non–COVID-19 cases with lung abnormalities due to other lung diseases will be correctly classified as not having COVID-19 but an alternative lung disease. It should be noted that the studies in our meta-analysis included patients between January and June 2020, a period with a high COVID-19 frequency (mean of 48.7%; range, 41.7%–59.8%). Specificity is likely to decrease with lower COVID-19 frequency and increasing frequency of other viral lung infections, such as influenza (29).

Our study had some limitations. First, the included studies used RT-PCR, which is an imperfect reference standard with a reported sensitivity of 89% (95% CI: 81, 94) (30). Sensitivity of RT-PCR appears to be lower in elderly patients (30), which may be due to sampling error in these patients who are more likely to have poorer performance status (25), Furthermore, vendor-specific effects and differences in the quality assurance process may affect the performance of RT-PCR (30). However, RT-PCR is still the recommended method to confirm current COVID-19 infection (31–33). Second, because of the relatively low number of included studies, we did not perform subgroup or meta-regression analyses to explain statistical heterogeneity between studies. Geographic differences, nonreported prevalence of other

lung diseases, interobserver variability in chest CT assessment, RT-PCR performance, and some methodologic quality issues may have been potential sources of heterogeneity. Note that interobserver agreement varied from substantial to almost perfect for the CO-RADS (7,19) and from moderate to substantial for the RSNA classification system (7,22).

In conclusion, COVID-19 infection frequency was higher in patients categorized with higher CO-RADS and RSNA classification categories.

Our data may be useful for deciding on the probability of COVID-19 based on chest CT (along with clinical information and RT-PCR).

Author contributions: Guarantors of integrity of entire study, R.M.K., T.C.K.; 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, all authors; statistical analysis, all authors; and manuscript editing, all authors.

Disclosures of Conflicts of Interest: R.M.K. disclosed no relevant relationships. **H.J.A.A.** disclosed no relevant relationships. **T.C.K.** disclosed no relevant relationships.

References

1. Johns Hopkins University School of Medicine. Coronavirus Resource Center. https://coronavirus.jhu.edu/. Updated June 13, 2020. Accessed August 24, 2020.
2. 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. Chest 2020;158(1):106–116.
3. Huang C, Wang Y, Li X, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet 2020;395(10223):497–506.
4. Adams HJA, Kwee TC, Yakar D, Hope MD, Kwee RM. Systematic Review and Meta-Analysis on the Value of Chest CT in the Diagnosis of Coronavirus Disease (COVID-19): Sol Scientiae, Illustra Nos. AJR Am J Roentgenol 2020;215(6):1342–1350.
5. Prokop M, van Everdingen W, van Rees Vellinga T, et al. CO-RADS: A Categorical CT Assessment Scheme for Patients Suspected of Having COVID-19-Definition and Evaluation. Radiology 2020;296(2):E97–E104.
6. 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 - Secondary Publication. J Thorac Imaging 2020;35(4):219–227.
7. de Jaegere TMH, Krdzalic J, Fasen BACM, Kwee RM. Radiological Society of North America Chest CT Classification System for Reporting COVID-19 Pneumonia: Interobserver Variability and Correlation with Reverse-Transcription Polymerase Chain Reaction. Radiol Cardiothorac Imaging 2020;2(3):e200213.
8. Suchá D, Van Hamersvelt RW, van den Hoven AF, de Jong PA, Verkooijen HM. Suboptimal Quality and High Risk of Bias in Diagnostic Test Accuracy Studies on Chest Radiography and Computed Tomography in the Acute Setting of the COVID-19 Pandemic: A Systematic Review. Radiol Cardiothorac Imaging 2020;2(4):e200342.
9. PRISMA. Transparent reporting of systematic reviews and meta-analyses. http://www.prisma-statement.org/. Updated 2015. Accessed August 24, 2020.
10. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med 2011;155(8):529–536.
11. Reitsma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. J Clin Epidemiol 2005;58(10):982–990.
12. OpenMeta[Analyst]. http://www.cebm.brown.edu/openmeta/. Accessed August 24, 2020.
13. The R Project for Statistical Computing. https://www.r-project.org/. Accessed August 24, 2020.
14. Mada-package. https://www.rdocumentation.org/packages/mada/versions/0.5.8/topics/mada-package. Accessed August 24, 2020.
15. Byrne D, O'Neill SB, Müller NL, et al. RSNA Expert Consensus Statement on Reporting Chest CT Findings Related to COVID-19: Interobserver Agreement Between Chest Radiologists. Can Assoc Radiol J 2020. 10.1177/0846537120938328. Published online July 2, 2020.
16. Lang M, Som A, Mendoza DP, et al. Detection of Unsuspected Coronavirus Disease 2019 Cases by Computed Tomography and Retrospective Implementation of the Radiological Society of North America/Society of Thoracic Radiology/American College of Radiology Consensus Guidelines. J Thorac Imaging 2020. 10.1097/RTI.0000000000000542. Published online June 17, 2020.
17. Adams SJ, Dennie C. Chest imaging in patients with suspected COVID-19. CMAJ 2020;192(25):E676.
18. Kwee RM, Krdzalic J, Fasen BACM, de Jaegere TMH; COVID-19 CT Investigators South-East Netherlands (CISEN) Study Group. CT Scanning in Suspected Stroke or Head Trauma: Is it Worth Going the Extra Mile and Including the Chest to Screen for COVID-19 Infection? AJNR Am J Neuroradiol 2020;41(7):1165–1169.
19. Fujioka T, Takahashi M, Mori M, et al. Evaluation of the Usefulness of CO-RADS for Chest CT in Patients Suspected of Having COVID-19. Diagnostics (Basel) 2020;10(9):E608.
20. De Smet K, De Smet D, Ryckaert T, 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.
21. Falaschi Z, Danna PSC, Arioli R, et al. Chest CT accuracy in diagnosing COVID-19 during the peak of the Italian epidemic: A retrospective correlation with RT-PCR testing and analysis of discordant cases. Eur J Radiol 2020;130:109192.
22. Ciccarese F, Coppola F, Spinelli D, et al. Diagnostic Accuracy of North America Expert Consensus Statement on Reporting CT Findings in Patients with Suspected COVID-19 Infection: An Italian Single Center Experience. Radiol Cardiothorac Imaging 2020;2(4):e200312.
23. Hermans JJR, Groen J, Zwets E, et al. Chest CT for triage during COVID-19 on the emergency department: myth or truth? Emerg Radiol 2020;27(6):641–651.
24. Korevaar DA, Kootte RS, Smits LP, et al. Added value of chest computed tomography in suspected COVID-19: an analysis of 239 patients. Eur Respir J 2020;56(2):2001377.
25. Miranda Magalhães Santos JM, Paula Alves Fonseca A, Pinheiro Zarattini Anastacio E, Formagio Minenelli F, Furtado de Albuquerque Cavalcanti C, Borges da Silva Teles G. Initial Results of the Use of a Standardized Diagnostic Criteria for Chest Computed Tomography Findings in Coronavirus Disease 2019. J Comput Assist Tomogr 2020;44(5):647–651.
26. Dofferhoff ASM, Swinkels A, Sprong T, et al. Diagnostic algorithm for COVID-19 at the ER [in Dutch]. Ned Tijdschr Geneeskd 2020;164:D502.
27. Adams HJA, Kwee TC, Yakar D, Hope MD, Kwee RM. Chest CT Imaging Signature of Coronavirus Disease 2019 Infection: In Pursuit of the Scientific Evidence. Chest 2020;158(5):1885–1895.
28. Adams HJA, Kwee TC, Kwee RM. Coronavirus Disease 2019 and Chest CT: Do Not Put the Sensitivity Value in the Isolation Room and Look Beyond the Numbers. Radiology 2020;297(1):E236–E237.
29. Altmayer S, Zanon M, Pacini GS, et al. Comparison of the computed tomography findings in COVID-19 and other viral pneumonia in immunocompetent adults: a systematic review and meta-analysis. Eur Radiol 2020;30(12):6485–6496.
30. Kim H, Hong H, Yoon SH. Diagnostic Performance of CT and Reverse Transcriptase Polymerase Chain Reaction for Coronavirus Disease 2019: A Meta-Analysis. Radiology 2020;296(3):E145–E155.
31. World Health Organization. Laboratory testing for 2019 novel coronavirus (2019-nCoV) in suspected human cases. https://www.who.int/publications/i/item/10665-331501. Updated March 19, 2020. Accessed August 24, 2020.
32. World Health Organization. Advice on the use of point-of-care immunodiagnostic tests for COVID-19. https://www.who.int/news-room/commentaries/detail/advice-on-the-use-of-point-of-care-immunodiagnostic-tests-for-covid-19. Updated April 8, 2020. Accessed August 24, 2020.
33. National Institute for Public Health and the Environment. Testing for COVID-19. https://www.rivm.nl/en/novel-coronavirus-covid-19/testing-for-covid-19. Updated August 17, 2020. Accessed August 24, 2020.