



Applicability of CO-RADS in an Anonymized Cohort Including Early and Advanced Stages of COVID-19 in Comparison to the Recommendations of the German Radiological Society and Radiological Society of North America

Anwendbarkeit der CO-RADS-Klassifikation in einer anonymisierten Kohorte mit frühen und fortgeschrittenen Krankheitsstadien im Vergleich zu den Empfehlungen der Deutschen Röntgengesellschaft und der Radiologischen Gesellschaft Nordamerikas

Authors

Birte Valentin¹ , Andrea Steuwe¹, Tobias Wienemann², Marcel Andree³, Verena Keitel⁴, Alexandra Ljimini¹ , Elisabeth Appel¹, Marie-Helen Köhler¹, Christin Rademacher¹, Joel Aissa¹, Gerald Antoch¹, Christina Loberg¹

Affiliations

- 1 Department of Diagnostic and Interventional Radiology, University Hospital Düsseldorf, Düsseldorf, Germany
- 2 Institute of Medical Microbiology and Hospital Hygiene, University Hospital Düsseldorf, Germany
- 3 Institute of Virology, University Hospital Düsseldorf, Düsseldorf, Germany
- 4 Department of Gastroenterology, Hepatology and Infectious Diseases, University Hospital Düsseldorf, Düsseldorf, Germany

Key words

SARS-CoV-2, pneumonia, viral, Multidetector Computed Tomography, Real-Time PCR, SARS-Associated Coronavirus

received 28.04.2021

accepted 03.01.2022

published online 24.02.2022

Bibliography

Fortschr Röntgenstr 2022; 194: 862–872

DOI 10.1055/a-1740-4310

ISSN 1438-9029

© 2022, Thieme. All rights reserved.

Georg Thieme Verlag KG, Rüdigerstraße 14,
70469 Stuttgart, Germany

Correspondence

Frau Dr. Birte Valentin
Department of Diagnostic and Interventional Radiology,
Universitätsklinikum Düsseldorf, Moorenstraße 5,
40225 Düsseldorf, Germany
Tel.: +49/211/8117752
Birte.Valentin@med.uni-duesseldorf.de

ZUSAMMENFASSUNG

Ziel Um eine einheitliche Befundung von Thorax-Computertomografien (CTs) mit Verdacht auf COVID-19 zu ermögli-

chen, wurden verschiedene Klassifikationen etabliert. CO-RADS klassifiziert Lungenparenchymveränderungen anhand ihrer Wahrscheinlichkeit für das Vorliegen einer SARS-CoV-2-Infektion. Diese Studie untersucht die retrospektive Anwendbarkeit der CO-RADS-Klassifikation in einer anonymisierten Kohorte im Vergleich zur DRG- und RSNA-Klassifikation. Verlaufsuntersuchungen wurden zusätzlich eingeschlossen, um ein fortgeschrittenes Krankheitsstadium zu simulieren. Als Referenzstandard dienen die Ergebnisse durchgeführter sequenzieller Reverse-Transkriptase-Polymerase-Kettenreaktionstests (qPCR).

Methoden Eingeschlossen wurden alle CT-Thorax Untersuchungen potenziell infizierter und nachweislich erkrankter Patienten zwischen dem 13. März und dem 30. November 2020. In diesem Zeitraum gab es aufgrund hoher Inzidenzwerte 2 regionale Lockdowns, wodurch eine hohe Vortestwahrscheinlichkeit vorliegt. Jede CT-Untersuchung wurde anonymisiert und anschließend nach CO-RADS-, DRG- oder RSNA-Klassifikation retrospektiv im Konsens durch 2 Radiologen (Assistenzarzt und Facharzt) ausgewertet. Die Befunde wurden mit den Ergebnissen der qPCR verglichen und eine statistische Auswertung wurde angefertigt.

Ergebnisse Insgesamt wurden 536 CT-Untersuchungen eingeschlossen. Die CO-RADS-, DRG- und RSNA-Klassifikationen erzielten einen negativ prädiktiven Wert von 96%/94%/95% (CO-RADS/DRG/RSNA), einen positiv prädiktiven Wert von 83%/80%/88%, eine Sensitivität von 86%/76%/80% und eine Spezifität von 96%/95%/97%. Die Prävalenz lag bei 20%.

Schlussfolgerung Alle Klassifikationen konnten verlässlich eine SARS-CoV-2-Infektion ausschließen. Nichtsdestotrotz lag eine hohe Vortestwahrscheinlichkeit bei unserem Studiensetting vor, die einen großen Einfluss auf die Klassifikationen hat. Daher bleibt es zu untersuchen, ob die Klassifikationen auch in Zukunft bei niedrigerer Prävalenz und Inzidenz von COVID-19 anwendbar sind.

Kernaussagen:

- Die CO-RADS-, DRG- und RSNA-Klassifikationen können helfen, Infizierte sicher in einer anonymisierten Kohorte zu erkennen
- Die Vortestwahrscheinlichkeit hat einen großen Einfluss auf die individuellen Klassifikationen
- Die Anonymisierung kann zu Fehlinterpretationen bei einem gravierenden Lungenbefall oder Residuen führen

ABSTRACT

Purpose Classifications were created to facilitate radiological evaluation of the novel coronavirus disease 2019 (COVID-19) on computed tomography (CT) images. The categorical CT assessment scheme (CO-RADS) categorizes lung parenchymal changes according to their likelihood of being caused by SARS-CoV-2 infection. This study investigates the diagnostic accuracy of diagnosing COVID-19 with CO-RADS compared to the Thoracic Imaging Section of the German Radiological Society (DRG) classification and Radiological Society of North America (RSNA) classification in an anonymized patient cohort. To mimic advanced disease stages, follow-up examinations were included as well.

Method This study includes all patients undergoing chest CT in the case of a suspected SARS-CoV-2 infection or an already confirmed infection between March 13 and November 30, 2020. During the study period, two regional lockdowns occurred due to high incidence values, increasing the pre-test probability of COVID-19. Anonymized CT images were reviewed retrospectively and in consensus by two radiologists applying CO-RADS, DRG, and RSNA classification. Afterwards, CT findings were compared to results of sequential real-time

reverse transcriptase polymerase chain reaction (qPCR) test performed during hospitalization to determine statistical analysis for diagnosing COVID-19.

Results 536 CT examinations were included. CO-RADS, DRG and RSNA achieved an NPV of 96%/94%/95% (CO-RADS/DRG/RSNA), PPV of 83%/80%/88%, sensitivity of 86%/76%/80%, and specificity of 96%/95%/97%. The disease prevalence was 20%.

Conclusion All applied classifications can reliably exclude a SARS-CoV-2 infection even in an anonymous setting. Nevertheless, pre-test probability was high in our study setting and has a great influence on the classifications. Therefore, the applicability of the individual classifications will become apparent in the future with lower prevalence and incidence of COVID-19.

Key Points:

- CO-RADS, DRG, and RSNA classifications help to reliably detect infected patients in an anonymized setting
- Pre-test probability has a great influence on the individual classifications
- Difficulties in an anonymized study setting are severe pulmonary changes and residuals.

Citation Format

- Valentin B, Steuwe A, Wienemann T et al. Applicability of CO-RADS in an Anonymized Cohort Including Early and Advanced Stages of COVID-19 in Comparison to the Recommendations of the German Radiological Society and Radiological Society of North America. *Fortschr Röntgenstr* 2022; 194: 862–872

Introduction

Since the first cases officially reported to the World Health Organization (WHO) in December 2019, the new coronavirus disease 2019 (COVID-19) has not been under control as a pandemic [1]. The gold standard of diagnostics is a real-time quantitative polymerase chain reaction (qPCR) test from nasopharyngeal and oropharyngeal swabs. Nevertheless, false-negative results may also be present with this method, depending on the time and method of testing [2, 3]. Low-dose chest computed tomography (CT) without contrast media has been recommended by the Fleischner Society and the Thoracic Imaging Section of the German Radiological Society (DRG) in cases of patients with a clinical presentation, medical history indicative of COVID-19, or aggravation of their symptoms [4, 5].

To facilitate radiological evaluation of this novel coronavirus on CT, the categorical CT assessment scheme CO-RADS was published by Prokop et al. in April 2020 [6]. CO-RADS rates suspected pulmonary involvement of COVID-19 using categories 1 (very low) to 5 (very high), depending on visible lung features. CO-RADS category 6 is assigned in the case of positive qPCR detection. Typical obligatory features of pulmonary involvement

of COVID-19 according to Prokop et al. are ground glass opacities, with or without consolidations, in lung regions close to visceral pleural surfaces and fissures, and multifocal bilateral distribution [6]. Confirmatory patterns are on the one hand unsharp demarcation or (half) rounded shape in ground glass regions, or on the other hand, sharp demarcation, outlining the shape of multiple adjacent secondary pulmonary lobules [6]. Further confirmatory patterns are crazy paving patterns compatible with organizing pneumonia and thickened vessels within parenchymal abnormalities found in all confirmatory patterns [6]. Besides this classification, there are other classifications such as the DRG or Radiological Society of North America (RSNA) classification (for comparison between classifications see Table [1, 7–9]).

However, the typical obligatory features of pulmonary involvement of COVID-19 are non-specific and may also appear in other types of viral pneumonia like influenza, cytomegalovirus, and adenovirus [10–13].

This retrospective study examines the CO-RADS classification in comparison to DRG and RSNA classifications in an anonymized cohort including follow-up examinations in clinical exacerbation, which allows the additional evaluation of their applicability in early and advanced disease stages.

Materials and methods

Ethics

The institutional review board approved this study and waived patient informed consent.

Patient cohort

This retrospective study includes patients that underwent a chest CT examination in the radiology department between March 13 and November 30, 2020. A further inclusion criterion was a suspected or confirmed SARS-CoV-2 infection, i. e., patients with new-onset respiratory symptoms, fever, loss of smell and/or taste, or confirmed contact with infected persons in the past 2 weeks. Routinely, patients presenting to the emergency department due to a pulmonary infection underwent a clinical examination, qPCR test, blood sampling and, if indicated (for indications see below), a chest CT. There were no exclusion criteria for this study.

Reference standard: qPCR

Patients were tested for suspected COVID-19 at the time of referral using a qPCR according to a standardized protocol of our virology department. The initial qPCR test (nasopharyngeal + oropharyngeal) was performed on admission to the hospital by trained and experienced emergency department medical staff. In the case of a positive CT result or an urgent suspicion of a SARS-CoV-2 infection despite initial negative qPCR result, a total of at least 2 qPCR tests were further performed.

Indications for CT and CT acquisition procedure

Since qPCR testing was considered the gold standard for evaluating SARS-CoV-2 infection, for radiation protection reasons, and according to the recommendations of the Thoracic Imaging Section of the German Radiological Society, chest CT was not performed as a screening method [14]. For each performed CT examination, justifying indication of using ionizing radiation for image acquisition was available. Depending on the clinical indication, patients received a chest CT either without or with contrast agent. Justifying indications were, for example, a massive worsening of the patient's symptoms with the need for oxygen, worsening of the ventilation situation, or suspected pulmonary artery embolism.

All patients with disease aggravation and urgent clinical suspicion of SARS-CoV-2 infection underwent non-enhanced low-dose chest CT (effective dose approximately 1 mSv for a person with a height of 170 cm and weight of 70 kg) in addition to a qPCR test. Contrast agents were applied only in cases with an additional differential-diagnostic question, e. g., pulmonary embolism (6/537 (1 %) examinations). Hospitalized infected patients or patients transferred from other hospitals during illness underwent CT if clinical exacerbation occurred during hospitalization.

All patients were examined on one of three state-of-the-art CT scanners (Somatom Definition Edge (scanner A), Somatom Definition Flash (scanner B), and Somatom Definition AS (scanner C) (all Siemens Healthineers, Forchheim, Germany). Patients were imaged in supine position with elevated arms and in breath-hold

technique following maximal inspiration. The scan range was defined from lung apex to base. The applied CT protocol parameters were as described in ► **Table 1**. Images were reconstructed iteratively using ADMIRE (scanner A) and SAFIRE (scanners B, C) (both Siemens Healthineers, Forchheim, Germany).

CT image evaluation

Each examination was anonymized and assigned a random number. Thus, follow-up examinations performed during hospitalization could not be related to the primary examination. Follow-up examinations were only included if there was a new suspicion, a continuing suspicion, or an aggravation of a confirmed SARS-CoV-2 infection. These follow-up examinations were included in order to test the respective classifications for detection at an early stage in case of new onset of symptoms during hospitalization or advanced stage of the disease in the case of aggravation of an infection. After anonymization, CT images were retrospectively reviewed in consensus by a resident and an experienced board-certified radiologist. Readers had no information on age, sex, previous diseases, laboratory, symptoms, or qPCR test results. CO-RADS classification (categories 1–5) and the recommendations of the DRG (categories 4–1) as well as of RSNA (categories 1–4) were applied [14, 15] (see ► **Table 1**). The German Radiological Society categorizes parenchymal changes according to CT changes: suggestive of COVID-19 pneumonia, indeterminate (COVID-19 pneumonia possible), suggestive of an alternative diagnosis (e. g., bacterial pathogen spectrum), and CT changes with no evidence of pneumonic opacities [14]. The RSNA classification also differentiates between typical, indeterminate, atypical parenchymal changes, and negative for pneumonia [9]. The occurrence of parenchymal changes (e. g., ground glass opacities, consolidations, crazy paving pattern, and thickened interlobular septa), their location (bilateral, peripheral, posterior, and basal emphasis), their manifestation (e. g., round or geographically configured), and other findings (e. g., pleural and pericardial effusion) were evaluated.

Radiation exposure

CT image data, volumetric CT dose index ($CTDI_{vol}$), dose length product (DLP), and scan length were collected in the local picture archive and communication system (SECTRA Medical, Sweden). Patient weight and height were documented. The effective dose was calculated using the tube potential-specific conversion factors published by Deak et al. ($k_{100kVp} = 0.0144 \text{ mSv}/(\text{mGy}\cdot\text{cm})$), $k_{120kVp} = 0.0145 \text{ mSv}/(\text{mGy}\cdot\text{cm})$), using ICRP publication 103 tissue weighting factors [16].

Data analysis

Statistical analysis was performed using SPSS version 22 (IBM, Chicago, IL, USA), GraphPad Prism 8.0 (GraphPad Software, Inc., San Diego, California, USA) and Microsoft Excel 2016 (Redmond, WA, USA). For continuous values, the mean and standard deviation with the corresponding ranges (minimum-maximum) are provided. A Mann-Whitney U-test for independent samples was performed to assess differences in CT findings in each group of findings. The level of significance was $p < 0.05$. ROC analysis was

► **Table 1** Overview of the CO-RADS, DRG, and RSNA classification with the associated parenchymal changes of the respective categories [7–9].

► **Tab. 1** Übersicht über die CO-RADS-, DRG- und RSNA-Klassifikation mit den zugehörigen Parenchymveränderungen der jeweiligen Kategorien.

Level of suspicion	CO-RADS (categories 1–5) [7]		DRG (categories 4–1) [8]		RSNA (categories 1–4) [9]	
Very low	1	no CT features suggest pneumonia	4	no CT features suggest pneumonia	1	no CT features suggest pneumonia
Low	2	tree in bud, centrilobular nodular patterns, consolidation, cavities, ground glass opacities location: centrilobular emphasis, lobar or segmental	3	noduli, tree in bud, peribronchial infiltrate, consolidations, caverns, bronchial wall thickening, mucus plugging, pleural effusion location: lobar or segmental	2	consolidation, noduli, tree in bud, cavitation, smooth interlobular septal thickening with pleural effusion location: isolated lobar or segmental
Unsure	3	ground glass opacities, crazy paving pattern, homogeneous location: peri-hilar	2	ground glass opacities, crazy paving-pattern, consolidations, no round shape or non-geographically location: central emphasis	3	ground glass opacities (very few small or diffuse) with or without consolidation, no round shape location: perihilar, peripheral sparing, unilateral, multifocal
High	4	ground glass opacities (with or without subpleural consolidations and air bronchogram), crazy paving pattern, reverse halo sign, arcade-like sign location: unilateral, peribronchovascular or superimposed with pre-existing lung changes	x	x	x	x
Very high	5	ground glass opacities (with or without subpleural consolidations and air bronchogram), crazy paving pattern, reversed halo sign, arcade-like sign location: near visceral pleural surfaces, multifocal and bilateral emphasis	1	ground glass opacities (round shaped or geographically), crazy paving pattern, consolidations, signs of organizing pneumonia, intralobular vasodilatation, no mediastinal lymphadenopathy location: peripheral, posterior, no subpleural sparing, bilateral, multifocal	4	ground glass opacities with or without consolidation, crazy paving pattern, multifocal, round shaped, reverse halo sign or other findings of organizing pneumonia location: peripheral, bilateral

x: no equivalent category exists, DRG: Deutsche Röntgengesellschaft, RSNA: Radiological Society of North America.

calculated using SPSS. We calculated the sensitivity, specificity, positive and negative predictive values, and diagnostic prevalence using qPCR results as the reference and calculated 95 % confidence intervals (CI) using GraphPad Prism and the epiR package in R [17].

Results

Patient population

The patient population included a total of 500 patients (63 % male and 37 % female) with a mean age of 69 ± 16 (range: 12–100) years and a mean body mass index (BMI) of 26.7 ± 5.8 (range: 12.8–49.0) kg/m² (► **Table 2**).

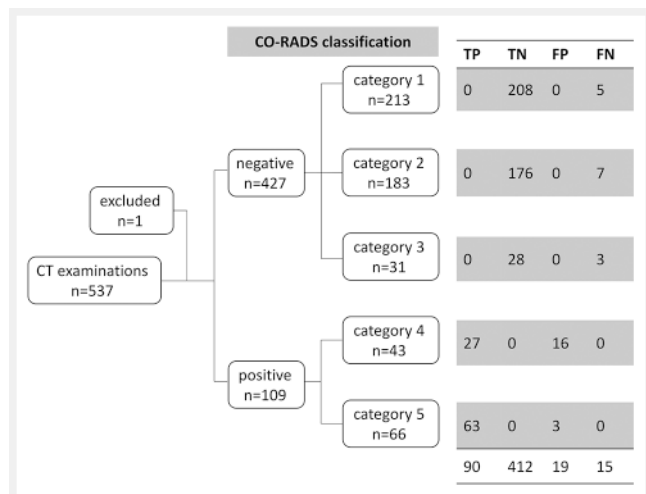
Including follow-up examinations, 537 chest CT examinations in 500 patients were evaluated in this study. In 94/500 (19 %) patients, active SARS-CoV-2 infection was detected by means of qPCR. In total, 106/537 (20 %) examinations were performed in qPCR-positive patients, while 431/537 (80 %) examinations were performed in qPCR-negative patients (see ► **Fig. 1**).

► **Table 2** Patient cohort information with CT-associated radiation exposure.

► **Tab. 2** Übersicht über das Patientenkollektiv und die mit den CT-Untersuchungen einhergehende Strahlenbelastung.

Parameter	Value
gender [m/f]	314/186
BMI [kg/m ²]	26.7 ± 5.8 (12.8–49)
CTDI _{vol} [mGy]	3.2 ± 1.3 (1.3–9.4)
DLP [mGycm]	102.9 ± 42.7 (41.8–336.0)
effective dose [mSv]	1.5 ± 0.6 (0.6–4.8)
scan length [cm]	30.0 ± 3.2 (21.0–38.9)

BMI: body mass index, FOV: field of view, DLP: dose length product, CTDI_{vol}: volumetric computed tomography dose index.



► **Fig. 1** Overview of the included computed tomography (CT) examinations and CO-RADS results. Abbreviations – TP: true positive, TN: true negative, FP: false positive, FN: false negative.

► **Abb. 1** Übersicht über die inkludierten Computertomografieuntersuchungen und erfolgten CO-RADS-Einteilungen. Abkürzungen – TP: richtig positiv, TN: richtig negativ, FP: falsch positiv, FN: falsch negativ.

Diagnostic performance of CO-RADS

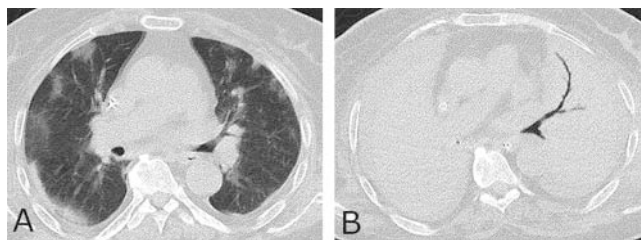
In total, 536/537 (99.8%) examinations were classifiable by CO-RADS. One CT examination was not classifiable due to severe lung parenchymal changes (only one bronchial tree was ventilated, see ► **Fig. 2**). For this reason, the examination was excluded.

Results of Fisher's exact tests to determine the separation category between potentially infected and potentially not infected are presented in ► **Table 3**. The best combination of negative predictive value (NPV) and positive predictive value (PPV) was obtained when categories 4 and 5 were considered potentially infected and categories 1–3 were considered potentially uninfected. Including category 3 among the potentially infected patients resulted in a slightly higher NPV at the expense of the PPV and specificity (for results see ► **Table 3**).

For the other 536 examinations, there were 90 true-positive, 412 true-negative, 19 false-positive, and 15 false-negative CO-RADS classifications (see ► **Fig. 1**). In total, 213 patients were classified CO-RADS 1, 183 patients CO-RADS 2, 31 patients CO-RADS 3, 43 patients CO-RADS 4, and 66 patients CO-RADS 5. The sensitivity was therefore 86% [95% CI: 78–92%], the specificity was 96% [95% CI: 93–97%], the PPV was 83% [95% CI: 74–89%], the NPV was 96% [95% CI: 94–98%], and the diagnostic accuracy was 94% [95% CI: 91–96%] (see ► **Table 3**).

Comparison of diagnostic performance between CO-RADS, DRG, and RSNA classification

The best combination of NPV and PPV for the DRG as well as the RSNA classification was obtained when category 1 (for DRG) and 4 (for RSNA) were considered potentially infected, and categories 2–4 (for DRG) and 1–3 (for RSNA) were considered potentially uninfected. Including category 2 (for DRG) and 3 (for RSNA) among



► **Fig. 2** A two-week course of disease in an infected patient is shown, each in the axial lung window. Image **A** was correctly classified as positive with visible peripheral ground glass opacities. Image **B** was excluded from the study due to severe acute respiratory distress syndrome without residual ventilated lung parenchyma. The patient was oxygenated via extracorporeal membrane oxygenation.

► **Abb. 2** Zweiwöchiger Verlauf von COVID-19 bei einem infizierten Patienten; gezeigt sind axiale Schichten im Lungenfenster. **A** Richtig positive Klassifizierung mittels CO-RADS anhand der sichtbaren peripheren Milchglastrübungen. **B** Von der Studie ausgeschlossene CT-Untersuchung nach Auftreten des schweren akuten respiratorischen Syndroms ohne verbleibendes ventiliertes Lungenparenchym.

the potentially infected patients resulted in a slightly higher NPV at the expense of the PPV and specificity (see ► **Table 3**).

Compared to the CO-RADS classification, both the DRG and the RSNA classification achieved a slightly lower sensitivity. The remaining parameters were similar (see ► **Table 3**).

The ROC analysis showed an AUC of 0.933 for CO-RADS, an AUC of 0.917 for DRG, and an AUC of 0.907 for the RSNA classification (see ► **Fig. 5a**).

CT findings

In 343/536 (64%) ground glass opacities, 304/536 (57%) consolidations, and 104/536 (19%) crazy paving patterns were visible. There was a bilateral manifestation in 281/536 (52%) of the cases. Pulmonary findings favored the peripheral parts of the lung in 210/536 (39%) of the cases and in 205/536 (38%) cases the manifestations emphasized the lower parts of the lung. The ROC analysis showed the highest AUC for peripheral emphasis (AUC 0.807), the number of affected lobes of the lung (AUC 0.807), bilateral emphasis (AUC 0.743), crazy paving pattern (AUC 0.735), ground glass opacities (AUC 0.682), posterior emphasis (AUC 0.672), consolidations (AUC 0.645), and thickened interlobular septa (AUC 0.621) (see ► **Fig. 5b, c**).

No findings indicating pneumonia were found in 56/536 (10%) examinations. All CT findings are shown in ► **Table 4**. Results of the statistical analysis can be found in ► **Table 5** and ► **Fig. 5**.

CT findings in positive classifications (CO-RADS 4–5)

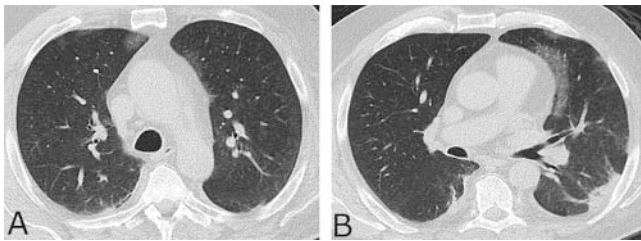
Ground glass opacities were visible in 97% of the examinations, consolidations in 82%, and crazy paving pattern in 56%. Bilateral manifestations were found in 94% of cases, emphasis of the lower lung lobes in 59%, emphasis of the peripheral lung lobes in 89%, and emphasis of the posterior lung lobes in 63%.

► **Table 3** Diagnostic performance of chest computed tomography by means of CO-RADS, DRG, and RSNA classification with quantitative polymerase chain reaction test result as a reference. Results of Fisher's exact test to determine the separation category between potentially infected and potentially not infected. The specified categories in the table are each considered potentially infected.

► **Tab. 3** Diagnostische Performance der Thoraxcomputertomografie mittels CO-RADS-, DRG- und RSNA-Klassifikation. Referenz ist die quantitative Polymerasekettenreaktion (qPCR). Ergebnisse des Fisher's exact test zur Bestimmung der Trennkategorie zwischen potenziell infizierten und potenziell nicht infizierten Patienten. Die in der Tabelle angegebenen Kategorien gelten jeweils als potenziell infiziert.

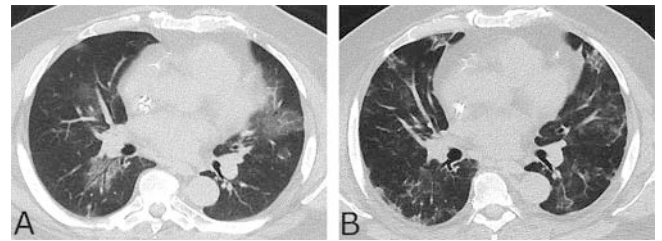
	CO-RADS (categories 4–5)	CO-RADS (categories 3–5)	DRG (category 1)	DRG (categories 1–2)	RSNA (category 4)	RSNA (categories 3–4)
Sensitivity (% [95% CI])	86 [78–92]	89 [81–93]	76 [67–83]	88 [80–93]	80 [71–86]	83 [75–89]
Specificity (% [95% CI])	96 [93–97]	89 [86–92]	95 [93–97]	89 [86–92]	97 [95–98]	94 [92–96]
Positive predictive value (% [95% CI])	83 [74–89]	66 [58–74]	80 [71–87]	67 [58–74]	88 [79–93]	78 [70–85]
Negative predictive value (% [95% CI])	96 [94–98]	97 [95–98]	94 [92–96]	97 [94–98]	95 [93–97]	96 [93–97]
Disease prevalence (%)	20 [17–24]					

CI: confidence interval.



► **Fig. 3** This is an example of a false-positive finding. Axial lung windows at different heights (A, B) are presented from one examination. Peripheral ground glass opacities, as well as paracardial crazy paving pattern are shown here. Nevertheless, the patient tested negative on three consecutive qPCR tests after the examination.

► **Abb. 3** Beispiel einer falsch positiven Klassifizierung. A, B) Gezeigt sind axiale Lungenschnitte in unterschiedlichen Höhen in einem Patienten. Hier traten periphere Milchglstrübungen und parakardiales crazy paving pattern auf. Nichtsdestotrotz waren 3 aufeinanderfolgende qPCR-Tests bei diesem Patienten negativ.



► **Fig. 4** Shown is the disease course of a patient in axial lung window during active infection A and follow-up examination with existing negative qPCR result B. In the follow-up examination, the previously affected regions continued to show ground glass opacities but also residual changes. Therefore, examination B was falsely classified as positive.

► **Abb. 4** Axiale Schichten der CT-Untersuchung eines infizierten Patienten zum Zeitpunkt der aktiven Infektion A und in einer Nachuntersuchung B mit negativem qPCR-Ergebnis. In der Nachuntersuchung zeigen sich sowohl weiterhin Milchglstrübungen in den betroffenen Regionen als auch verbleibende Lungenveränderungen.

CT findings in true-positive classifications

Ground glass opacities were visible in 97% of the examinations, consolidations in 82%, and crazy paving pattern in 59%. Bilateral manifestations were found in 93% of cases, emphasis of the lower lung lobes in 59%, emphasis of the peripheral lung lobes in 94%, and emphasis of the posterior lung lobes in 67%.

CT findings in false-positive classifications:

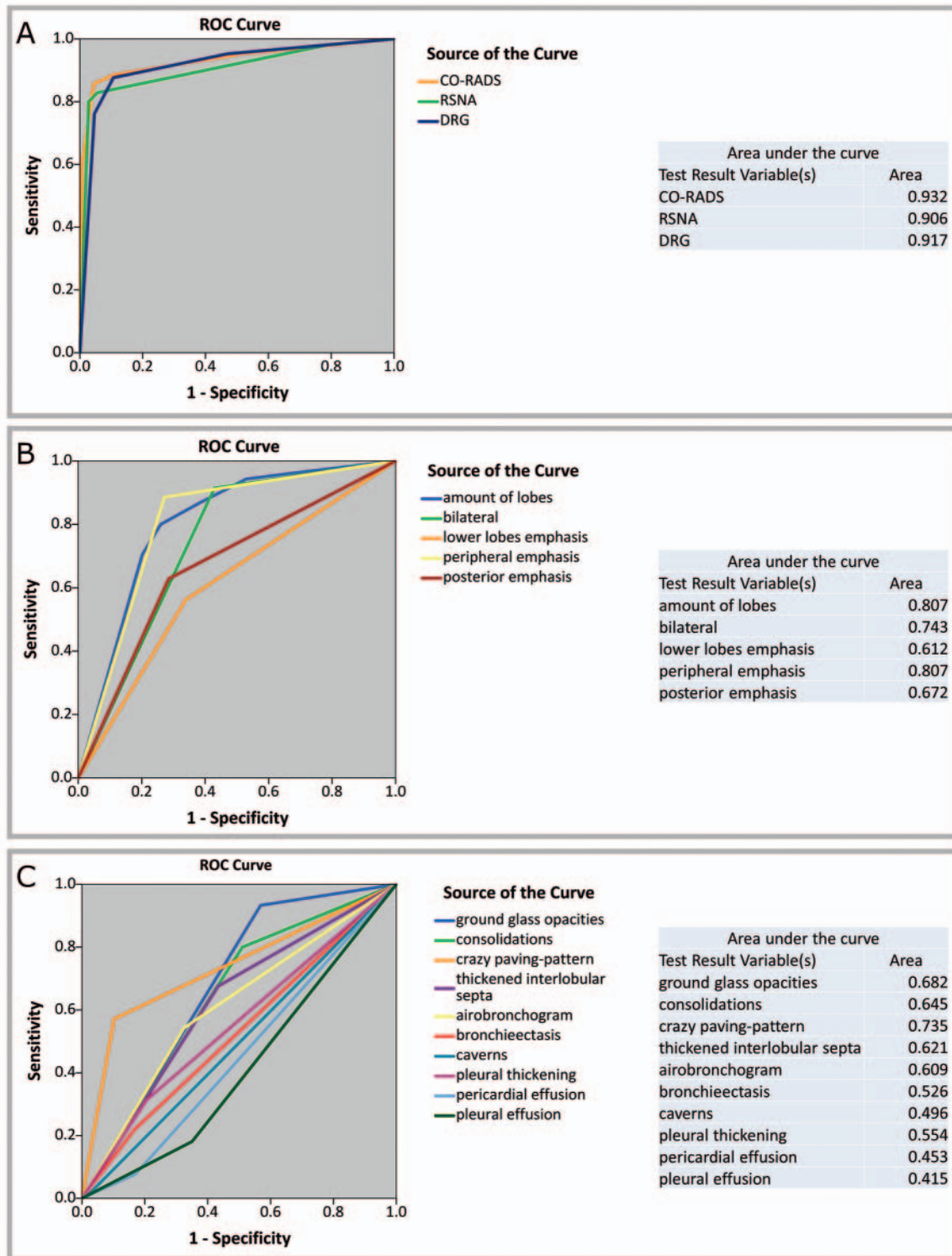
Ground glass opacities were visible in 100% of the examinations, consolidations in 79%, and crazy paving pattern in 42%. Bilateral manifestations were found in 95% of cases, emphasis of the lower

lung lobes in 58%, emphasis of the peripheral lung lobes in 63%, and emphasis of the posterior lung lobes in 47% (see ► Fig. 3, 4).

In some patients classified as false positive, other viruses, such as cytomegalovirus, could be detected.

CT findings in negative classifications (CO-RADS 1–3)

Ground glass opacities were visible in 55% of the examinations, consolidations in 50%, and crazy paving pattern in 10%. Bilateral manifestations were found in 42% of cases, emphasis of the lower



► **Fig. 5** Results of the ROC analysis of qPCR-positive and negative patients and all classifications. CO-RADS achieved slightly higher AUC compared to RSNA and DRG. Best area under the curve values were obtained with peripheral emphasis, amount of lobes, bilateral emphasis, and crazy paving pattern.

► **Abb. 5** Dargestellt sind die ROC-Analysen der qPCR-positiven und -negativen Patienten und der Klassifikationen. CO-RADS erzielte eine geringfügig höhere AUC im Vergleich zu der RSNA- und DRG-Klassifikation. Die höchsten AUC-Werte erzielten die periphere Betonung, die Anzahl der betroffenen Lungenlappen, der bilaterale Lungenbefall und das crazy paving pattern.

► **Table 4** Manifestations seen on computed tomography images, depending on CO-RADS classification and qPCR results.

► **Tab. 4** Lungenbefunde des Patientenkollektivs in Abhängigkeit der CO-RADS-Klassifikation und qPCR-Ergebnissen.

Manifestation	CO-RADS				qPCR	
	Positive score 4–5 n = 109	Negative score 1–3 n = 427	TP n = 90	FP n = 19	Positive n = 106	Negative n = 431
Ground glass opacities (%)	97	55	97	100	93	57
Consolidation (%)	82	50	82	79	81	51
Crazy paving pattern (%)	56	10	59	42	58	10
Thickened interlobular septa (%)	66	44	69	53	68	43
Air bronchogram (%)	56	32	54	63	54	32
Caverns (%)	1	3	1	0	2	3
Pleural thickening (%)	28	22	31	16	32	21
Pneumothorax (%)	1	0	1	0	2	0
Number of lobes (n)						
▪ 0	0	141	0	0	4	139
▪ 1	0	64	0	0	2	64
▪ 2	6	66	5	1	7	65
▪ 3	16	47	10	6	7	52
▪ 4	11	25	9	2	11	25
▪ 5	76	84	66	10	74	86
Bilateral infestation (%)	94	42	93	95	92	43
Emphasis lower lobes (%)	59	33	59	58	58	34
Emphasis periphery (%)	89	27	94	63	90	27
Emphasis posterior (%)	63	28	67	47	63	29
Pleural effusion (%)	16	36	16	16	18	35
Pericardial effusion (%)	6	17	7	5	8	17

TP: true positive, FP: false positive, qPCR: quantitative polymerase chain reaction.

lung lobes in 33 %, emphasis of the peripheral lung lobes in 27 %, and emphasis of the posterior lung lobes in 28 %.

CT findings in true-negative classifications

Ground glass opacities were visible in 55 % of the examinations, consolidations in 50 %, and crazy paving pattern in 9 %. Bilateral manifestations were found in 41 % of cases, emphasis of the lower lung lobes in 33 %, emphasis of the peripheral lung lobes in 25 %, and emphasis of the posterior lung lobes in 28 %.

CT findings in false-negative classifications

Ground glass opacities were visible in 73 % of the examinations, consolidations in 67 %, and crazy paving pattern in 47 %. Bilateral manifestations were found in 80 % of cases, emphasis of the lower lung lobes in 47 %, emphasis of the peripheral lung lobes in 60 %, and emphasis of the posterior lung lobes in 47 %.

CT findings in CO-RADS true positive vs. CO-RADS true negative

Differences occurred in parenchymal changes such as ground glass opacities (97 %:55 %, $p < 0.05$), consolidation (82 %:50 %, $p < 0.05$), crazy paving pattern (59 %:9 %, $p < 0.05$), thickened interlobular septa (69 %:43 %, $p < 0.05$). Bilateral manifestation (93 %:41 %, $p < 0.05$) and emphasis of the lower (59 %:33 %, $p < 0.05$), peripheral (94 %:25 %, $p < 0.05$) and posterior (67 %:28 %, $p < 0.05$) lung lobes were more common in categories 4–5. Pleural (16 %:36 %, $p < 0.05$) and pericardial (7 %:17 %, $p < 0.05$) effusion were more often associated with CO-RADS categories 1–3 (see ► **Table 5**).

CT findings in qPCR positive vs. qPCR negative

Significant differences occurred in parenchymal changes such as ground glass opacities (92 %:57 %, $p < 0.05$), consolidation (81 %:51 %, $p < 0.05$), crazy paving pattern (58 %:10 %, $p < 0.05$),

► **Table 5** Results of the statistical comparison of the manifestations between CO-RADS-positive and qPCR-positive patients and between CO-RADS true-positive and false-positive patients. A p-value of <0.05 was considered significant.

► **Tab. 5** Ergebnisse der statistischen Vergleiche der Lungenbefunde zwischen CO-RADS-positiven und qPCR-positiven Patienten und zwischen CO-RADS richtig positiven und falsch positiven Patienten.

Manifestation	Statistical comparison				
	CO-RADS neg:pos	CO-RADS TN:TP	qPCR neg:pos	CO-RADS pos: qPCR pos	CO-RADS TP:FP
Ground glass opacities (%)	*	*	*	ns	ns
Consolidation (%)	*	*	*	ns	ns
Crazy paving pattern (%)	*	*	*	ns	ns
Thickened interlobular septa (%)	*	*	*	ns	ns
Air bronchogram (%)	*	*	*	ns	ns
Caverns (%)	ns	ns	ns	ns	ns
Pleural thickening (%)	ns	*	ns	ns	ns
Pneumothorax (%)	ns	*	*	ns	ns
Bilateral infestation (%)	*	*	*	ns	ns
Emphasis lower lobes (%)	*	*	*	ns	ns
Emphasis periphery (%)	*	*	*	ns	*
Emphasis posterior (%)	*	*	*	ns	ns
Pleural effusion (%)	*	*	*	ns	ns
Pericardial effusion (%)	*	*	*	ns	ns

TP: true positive, TN: true negative, qPCR: quantitative polymerase-chain reaction, pos: positive, neg: negative, *: significant difference, ns: no statistical difference.

and thickened interlobular septa (68%:43 %, $p < 0.05$). Bilateral manifestation (92%:43 %, $p < 0.05$) and emphasis of the lower (58%:34 %, $p < 0.05$), peripheral (98%:27 %, $p < 0.05$) and posterior (62%:29 %, $p < 0.05$) lung lobes were more common in qPCR-positive patients. Pleural (19%:35 %, $p < 0.05$) and pericardial (8%:17 %, $p < 0.05$) effusion were more often associated with qPCR-negative patients (see ► **Table 5**).

CT exposure parameters

For the CT examinations, the mean $CTDI_{vol}$ and DLP were 3.2 ± 1.3 (1.3–9.4) mGy and 102.9 ± 43.7 (41.8–336.0) mGy·cm, respectively, resulting in an effective dose of 1.5 ± 0.6 (0.6–4.8) mSv (see ► **Table 2**).

Discussion

Even in an anonymized study setting, it is possible to obtain satisfactory results when applying CO-RADS, DRG, and RSNA classifications to patients with suspected SARS-CoV-2 infection. Advanced stages of the disease are also well recognized, with exceptions such as severe ARDS.

Our data indicate similar sensitivities (CO-RADS: 86 %) to Schalekamp et al. (86 %), Fujioka et al. (87.8 %), and Smet et al. (32–85 %) and a higher sensitivity compared to Bellini et al. (61 %),

when CO-RADS classification is used [18–21]. Furthermore, we report a higher specificity (CO-RADS: 96 %) compared to Fujioka et al. (66.4 %), Bellini et al. (81 %), Schalekamp et al. (81 %), and similar specificity compared to Smet et al. (85–95 %) [18–21]. As performed in this study, patient information was blinded in the studies of Fujioka et al., Bellini et al., and Smet et al. [19–21]. Schalekamp et al. used patient information but were blinded to the qPCR test results [18]. One reason for our higher sensitivity and specificity compared to Bellini et al. could be their high number of readers (12 vs. 2) and non-consensual decision-making [21]. Additionally, the patient collective is important. Only patients with suspected or confirmed infection were included in the presented study. Thus, a cohort preselection was performed. A similar preselection was performed by Bellini et al., who included only patients with clinical suspicion of COVID-19 [21]. Despite the higher prevalence (25 %), Bellini et al. achieved poorer sensitivity and specificity. In comparison to the literature, we additionally obtained a higher NPV (CO-RADS: 96 %, DRG 94 %, RSNA 95 %) than Bellini et al. (CO-RADS: 77.4–86.7 %) and Smet et al. (CO-RADS: 90.3 %) [20, 21]. Hence, the assessed classification helps to reliably exclude COVID-19. Compared to the anonymized application of the DRG and RSNA classification, the CO-RADS classification performed slightly better with respect to sensitivity and NPV. This could be due to the study design, as the CO-RADS classification exclusively analyzes the lung parenchyma [6]. The

DRG and RSNA classifications additionally evaluate preexisting cardiac diseases in addition to the lung parenchyma. Due to anonymization, preexisting cardiac diseases could not be included in the evaluation. This could have an influential effect on the results obtained here for the respective classifications.

We observed significantly more thickened interlobular septa in infected patients in our cohort compared to non-infected patients. Thickened interlobular septa could occur due to an interstitial inflammatory response as well as a possible cardiac involvement of the infection. The presence of, e. g., pulmonary hypertension or cardiopulmonary congestion can result in the occurrence of ground glass opacities and thickened interlobular septa, thereby imitating crazy paving pattern and thus complicating the differentiation between COVID-19-induced parenchymal changes and additive changes to preexisting parenchymal changes. Nevertheless, thickened interlobular septa occurred more frequently in infected patients in our cohort. It is well known that viral infections can also affect the myocardium and lead to myocarditis [22]. Unfortunately, we did not have data on cardiac examinations, but this observation should be further monitored.

In addition, the radiologist must not rely solely on the presence of patterns, such as ground glass opacities. This is because they also occur in a variety of other viral diseases or systemic diseases. Ground glass opacities occurred in over 50 % of negatively classified patients and up to 100 % of false-positive patients. In this context, the specific appearance and location of the ground glass opacities is important: We observed that the presence of discrete ubiquitous changes or unilateral changes, usually localized in only one lobe, made infection less likely. However, if round-shaped ground glass opacities as well as a bilateral, multilobular, posterior, and peripheral occurrence are present, an infection is more likely [23–25]. These observations are consistent with the classifications.

One advantage of the CO-RADS classification is the higher number of subcategories. This allows refined categorization of patients, which may result in a lower number of patients in need of isolation. Nevertheless, it is important to be aware that some patients have only subtle or no parenchymal changes in the early stages of disease and can thus be easily misclassified.

All applied classifications have an intermediate category (CO-RADS 3, DRG 2, and RSNA 3) where the suspicion of an existing infection is uncertain. The number of patients in intermediate categories is influenced by the incidence and prevalence of COVID-19, other viral pneumonia, as well as underlying diseases or therapy-associated pulmonary alterations, such as those caused by drug toxicity or radiation. The number of patients in the intermediate categories increases as the corresponding prevalence and incidence of other viral diseases, underlying diseases, and therapy-associated pulmonary alterations increases. This is seasonal, especially in the case of viral diseases. Yet, there were verifiably fewer other viral illnesses in 2020, e. g., due to mouth-nose protection (MNP) [26]. Thus, we assume that the number of patients in the intermediate categories will decrease in case of continuing MNP in the autumn and winter months. The reasons for assigning SARS-CoV-2-infected patients to the intermediate category could be related to the timing of imaging at

the disease stage of a SARS-CoV-2 infection as well as individual parenchymal involvement.

In this study, a cut-off category of 4 was determined for all assessed classifications based on statistical analyses. Still, we cannot classify patients in the intermediate categories as non-infected in general. Especially in our study setting, there were variations in the prevalence of COVID-19 (including 2 lockdowns) and other viral diseases over the study period. Therefore, this study has an increased pre-test probability. An improvement of the assessment of the intermediate categories is possible, especially in a non-anonymized setting with knowledge of previous diseases, contact persons, origin of the patient (e. g., from a county region, a district with a higher prevalence of COVID-19, or traveler returning from a high prevalence region) as well as current therapies (e. g., chemotherapy, radiation of the thorax). These additive clinical parameters could further help with categorizing the level of suspicion.

In our opinion, all classifications are helpful and applicable in reporting. Classifications enable better comparability independent of the level of knowledge of the radiologist evaluating the CT images. Nevertheless, pre-test probability plays a major role in the application of classifications. Thus, the use of classifications when incidence is low could pose a risk of misclassifying healthy individuals as infected.

By choosing an anonymized study approach, we tried to minimize the influence of a preexisting qPCR test result and clinical information. Yet, some patients in our cohort with positive qPCR test results had only few, atypical, or no parenchymal changes at all, resulting in a false-negative CT diagnosis and reduced sensitivity. In these patients, COVID-19 could be present at a very early stage, meaning that parenchymal changes cannot be detected or are atypical at the time of CT. Another reason could be the individually lower immune response in the lung parenchyma of these patients, resulting in only mild parenchymal changes.

Pronounced pulmonary involvement in advanced disease could be a reason for misinterpretation. Especially in cases of ARDS, the lung parenchyma is severely altered bilaterally, that is why typical patterns of infection could no longer be reliably differentiated. In addition, all classifications do not consider the fate of patterns or parenchymal residuals after survived infection, resulting in reduced specificity [27]. However, since reinfections occur less frequently than primary infections, this issue is negligible when applying classifications. We assume that beyond this anonymized study setting, information about survived infection is available and thus the rate of false-positive findings due to residual changes is low. Nevertheless, there are patients who have gone through a mild infection and may not be aware of it. These patients could have residual changes as well, which could be scored as false positives.

The study interval (March to November 2020) included two waves of the pandemic, with improved qPCR tests (both time and sensitivity) at the time of the second wave. This could affect the number of CT examinations in a screening setting. However, the strict indication for CT in our hospital (CT was not applied as screening/triage method in patients) had less of an impact on the decision for CT than the indication for CT itself.

Although anamnestic information is useful, it can also be misleading regarding the estimation of the likelihood of a SARS-CoV-2 infection in the case of inconclusive CT findings. In this case, readers are probably more willing to classify a patient as infectious when contact to infected persons is confirmed. Nevertheless, potential infection chains are interrupted at the expense of a higher workload for the inpatient sector.

Conclusion

All applied classifications can reliably exclude a SARS-CoV-2 infection even in an anonymized setting. CO-RADS achieved slightly better results in our cohort than the other classifications. CO-RADS is suitable for initial assessment at disease onset but has limitations in advanced disease and post-inflammatory pulmonary residuals.

Clinical information, e. g., confirmed contact to infected persons, living in high-prevalence regions, or returning traveler, should be additional criteria to the classifications. In our opinion, this could increase diagnostic accuracy.

Nevertheless, the pre-test probability has a great influence on the classifications. Therefore, the applicability of the individual classifications will become apparent in the future with lower prevalence and incidence of COVID-19.

Conflict of Interest

The authors declare that they have no conflict of interest.

References

- [1] World Health Organization. WHO Coronavirus Disease Dashboard. Geneva, 2020
- [2] Chen Z, Li Y, Wu B et al. A Patient with COVID-19 Presenting a False-Negative Reverse Transcriptase Polymerase Chain Reaction Result. *Korean J Radiol* 2020; 21: 623–624
- [3] Li D, Wang D, Dong J et al. False-Negative Results of Real-Time Reverse-Transcriptase Polymerase Chain Reaction for Severe Acute Respiratory Syndrome Coronavirus 2: Role of Deep-Learning-Based CT Diagnosis and Insights from Two Cases. *Korean J Radiol* 2020; 21: 505–508
- [4] Rubin GD, Ryerson CJ, Haramati LB et al. The Role of Chest Imaging in Patient Management During the COVID-19 Pandemic. *Chest* 2020; 158: 106–116
- [5] Vogel-Claussen J, Ley-Zaporozhan J, Agarwal P et al. Recommendations of the Thoracic Imaging Section of the German Radiological Society for clinical application of chest imaging and structured CT reporting in the COVID-19 pandemic. *RoFo Fortschritte Auf Dem Gebiet Der Röntgenstrahlen Und Der Bildgeb Verfahren* 2020; 192: 633–640
- [6] 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: E97–E104
- [7] Penha D, Pinto EG, Matos F et al. CO-RADS: Coronavirus Classification Review. *J Clin Imaging Sci* 2021; 11: 9
- [8] AG Thoraxdiagnostik in der Deutschen Röntgengesellschaft. Kategorisierung der CT-Veränderungen bei COVID-19 Epidemie. 2020
- [9] Simpson S, Kay FU, Abbara S 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: e200152
- [10] Xie S, Lei Z, Chen X et al. Chest CT-based differential diagnosis of 28 patients with suspected corona virus disease 2019 (COVID-19). *Br J Radiol* 2020; 93: 20200243
- [11] Xie C, Tsakok M, Channon-Wells S et al. COVID-19 pneumonia and the masquerades. *BJR | case Reports* 2020; 6: 20200067
- [12] Li X, Fang X, Bian Y et al. Comparison of chest CT findings between COVID-19 pneumonia and other types of viral pneumonia: a two-center retrospective study. *Eur Radiol* 2020; 30: 5470–5478
- [13] Koo HJ, Lim S, Choe J et al. Radiographic and CT features of viral pneumonia. *RadioGraphics* 2018; 38: 719–739
- [14] Vogel-Claussen J, Ley-Zaporozhan J, Agarwal P et al. Recommendations of the Thoracic Imaging Section of the German Radiological Society for clinical application of chest imaging and structured CT reporting in the COVID-19 pandemic. *Fortschr Röntgenstr* 2020; 192: 633–640
- [15] Simpson S, Kay FU, Abbara S 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: e200152
- [16] Deak PD, Smal Y, Kalender WA. Multisection CT Protocols: Sex- and Age-specific Conversion Factors Used to Determine Effective Dose from Dose-Length Product. *Radiology* 2010; 257: 158–166
- [17] R Core Team. R: A Language and Environment for Statistical Computing. 2020
- [18] Schalekamp S, Bleeker-Rovers CP, Beenen LFM et al. Chest CT in the Emergency Department for Diagnosis of COVID-19 Pneumonia: Dutch Experience. *Radiology* 2020; 78: 203465
- [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* 2020; 10: 608
- [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: E30–E37
- [21] Bellini D, Panvini N, Rengo M et al. Diagnostic accuracy and interobserver variability of CO-RADS in patients with suspected coronavirus disease-2019: a multireader validation study. *Eur Radiol* 2021; 31: 1932–1940
- [22] Al-Akchar M, Kiel J. Acute Myocarditis. 2021
- [23] Hani C, Trieu NH, Saab I et al. COVID-19 pneumonia: A review of typical CT findings and differential diagnosis. *Diagn Interv Imaging* 2020; 101: 263–268
- [24] Carotti M, Salaffi F, Sarzi-Puttini P et al. Chest CT features of coronavirus disease 2019 (COVID-19) pneumonia: key points for radiologists. *Radiol Med* 2020; 125: 636–646
- [25] Schmitt W, Marchiori E. COVID-19: Round and oval areas of ground-glass opacity. *Pulmonology* 2020; 26: 246–247
- [26] Hills T, Kearns N, Kearns C et al. Influenza control during the COVID-19 pandemic. *Lancet* 2020; 396: 1633–1634
- [27] Han X, Fan Y, Alwalid O et al. Six-month Follow-up Chest CT Findings after Severe COVID-19 Pneumonia. *Radiology* 2021; 299: E177–E186