RSNA Expert Consensus Statement on Reporting Chest CT Findings Related to COVID-19: Interobserver Agreement Between Chest Radiologists

Danielle Byrne, FFR (RCSI)¹,², Siobhan B. O’Neill, FFR (RCSI)¹,², Nestor L. Müller, MD, PhD¹,², C. Isabela Silva Müller, MD, PhD³, John P. Walsh, FFR (RCSI)¹,², Sabeena Jalal, MBBS, MSc¹, William Parker, MD, BSSc¹,², Ana-Maria Bilawich, MD¹,², and Savvas Nicolaou, MD¹,²

Abstract

Purpose: To assess the interobserver variability between chest radiologists in the interpretation of the Radiological Society of North America (RSNA) expert consensus statement reporting guidelines in patients with suspected coronavirus disease 2019 (COVID-19) pneumonia in a setting with limited reverse transcription polymerase chain reaction testing availability. Methods: Chest computed tomography (CT) studies in 303 consecutive patients with suspected COVID-19 were reviewed by 3 fellowship-trained chest radiologists. Cases were assigned an impression of typical, indeterminate, atypical, or negative for COVID-19 pneumonia according to the RSNA expert consensus statement reporting guidelines, and interobserver analysis was performed. Objective CT features associated with COVID-19 pneumonia and distribution of findings were recorded. Results: The Fleiss kappa for all observers was almost perfect for typical (0.815), atypical (0.806), and negative (0.962) COVID-19 appearances (P < .0001) and substantial (0.636) for indeterminate COVID-19 appearance (P < .0001). Using Cramer V analysis, there were very strong correlations between all radiologists’ interpretations, statistically significant for all (typical, indeterminate, atypical, and negative) COVID-19 appearances (P < .001). Objective CT imaging findings were recorded in similar percentages of typical cases by all observers. Conclusion: The RSNA expert consensus statement on reporting chest CT findings related to COVID-19 demonstrates substantial to almost perfect interobserver agreement among chest radiologists in a relatively large cohort of patients with clinically suspected COVID-19. It therefore serves as a reliable reference framework for radiologists to accurately communicate their level of suspicion based on the presence of evidence-based objective findings.

Résumé

Objectif : Evaluator la variabilité inter-observateurs entre radiologistes thoraciques concernant les lignes directrices de la déclaration de consensus d’experts de la Radiological Society of North America (RSNA) pour le signalement des patients ayant une pneumonie COVID-19 suspectée dans un établissement où la disponibilité des tests de dépistage par rt-PCR est limitée. Méthodes : Des études de tomodensitométries (TDM) thoraciques effectuées chez 303 patients consécutifs suspects de COVID-19 ont été analysées par 3 radiologistes thoraciques formés dans le cadre d’un fellowship. Chaque cas a été classé selon sa forme (typique, indéterminé, atypique ou négatif) pour la pneumonie COVID-19 en suivant les lignes directrices pour leur signalement tirées de la déclaration de consensus d’experts de la RSNA ; une analyse inter-observateurs a ensuite été effectuée. Les caractéristiques objectives des TDM associées à la pneumonie COVID-19 et la répartition des constatations ont été consignées. Résultats : Le kappa de Fleiss pour tous les observateurs a été presque parfait pour toutes les formes typiques (0,815), atypiques

1 Department of Radiology, Vancouver General Hospital, British Columbia, Canada
2 University of British Columbia, Vancouver, British Columbia, Canada
3 Department of Radiology, Delfin Clinic, Salvador, Bahia, Brazil

Corresponding Author:
Danielle Byrne, FFR (RCSI), Department of Radiology, Vancouver General Hospital, Vancouver, British Columbia, Canada V5Z 1M9.
Email: dbyrne.radiology@gmail.com
Introduction

Coronavirus disease 2019 (COVID-19) is a novel infectious disease caused by severe acute respiratory syndrome coronavirus 2. It was first identified in the human population in December 2019 and spread rapidly around the world reaching pandemic status in March 2020. Reverse transcription polymerase chain reaction (RT-PCR) of respiratory specimens is the most widely used method for diagnosing COVID-19. However, in many clinical settings, RT-PCR testing may not be readily available, be limited to hospitalized patients, or may be initially falsely negative. Furthermore, patients with mild respiratory symptoms may be reluctant to travel to a PCR testing site or hospital.

Although most radiological societies and professional organizations have recommended against performing chest computed tomography (CT) in the workup of patients with suspected COVID-19, in resource-poor countries where RT-PCR is not widely available, a tentative diagnosis or exclusion is often based on clinical and CT findings alone. Additionally, several retrospective studies have shown that CT has greater sensitivity (86%-98%) and lower false-negative rate than RT-PCR. The sensitivity of nasopharyngeal swabs ranges from 42% to 71%, with false negatives observed more frequently early in the course of the disease.

This has led to the widespread use of chest CT in the workup of patients with suspected COVID-19 particularly in the setting of negative RT-PCR and high clinical suspicion. COVID-19 results in a spectrum of chest CT manifestations, which evolve over time ranging from peripheral predominant ground-glass opacities to an organizing pneumonia reaction pattern, with additional findings including crazy-paving and more diffuse ground glass, thus overlapping with imaging features of a variety of other disease processes including other infections, drug reaction, and inhalational exposure. As there are a number of etiologies with imaging findings that overlap with those of COVID-19, the inclusion of COVID-19 in the differential diagnosis of the radiology report may lead to unwarranted anxiety among referring physicians and patients. Optimal interpretation and reporting of the chest CT and clear, unambiguous communication with the referring physician are essential for optimal patient care as well as community and health care worker safety during the COVID-19 pandemic.

An 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 the Radiological Society of North America (RSNA), was published to assist radiologists in recognizing and describing lung imaging findings in a standardized manner in patients under investigation for COVID-19 pneumonia and provide clarity in communication with other health care providers. The purpose of this study was to assess the interobserver variability between chest radiologists in the evaluation of CT scans in a cohort of patients with respiratory symptoms and suspected COVID-19 pneumonia using the RSNA expert consensus reporting guidelines in a setting with limited RT-PCR testing availability.

Methods and Materials

Study Population

Approval was obtained from the institutional review board, and the need for informed consent was waived for this retrospective study. A total of 303 consecutive patients (158 males and 145 females) with a median age of 49 years (range, 18-96) who presented with respiratory symptoms and suspected COVID-19 pneumonitis were included. Due to the resource-limited nature of the hospital, reliable RT-PCR data were not available.

Imaging Protocol

Chest CT studies were acquired on a 32-slice single-source CT scanner (Siemens Somatom Scope) using a standardized CT technique: 110 kVp, 345 mA max, pitch of 1.5, 0.6 seconds rotation time, and 1 mm scan thickness. Inspiratory phase CT chest examinations were acquired with patients in the supine position without administration of intravenous contrast.

Imaging Analysis

Computed tomography images were extracted from the picture archiving and communication system, anonymized, and...
imported onto a secure browser-based viewing system for CT scans (MD.ai, available at https://www.md.ai/). The software displayed cross-sectional CT images only with soft tissue and lung kernel and allowed for dynamic scrolling, window width–window level adjustment, panning, and zoom. Three fellowship-trained chest radiologists, each with >10 years experience reading chest CT examinations (S.B.O.N., A.-M.B., and C.I.S.M.) performed diagnostic interpretation on all included CT studies. Observers, blinded to all patient information including RT-PCR result, recorded the presence or absence and distribution of typical COVID-19 imaging features (see Table 1) and scored each study according to the RSNA expert consensus reporting guidelines as to whether a study was typical, indeterminate, atypical, or negative for the presence of COVID-19 pneumonia.10

Statistical Analysis

Statistical analysis was performed using SPSS statistics version 25 (IBM). To quantify interobserver agreement, Cramer V and Fleiss kappa were determined across observers. Kappa values were obtained by reporting scores of each observer (according to the RSNA expert consensus statement on reporting chest CT findings related to COVID-19) to the median of the other 2 observers. Interobserver agreement was considered slight for a kappa value of 0.01 to 0.20, fair for 0.21 to 0.40, moderate for 0.41 to 0.60, substantial for 0.61 to 0.80, and almost perfect for 0.81 to 1.00.17 The result of a Cramer’s V-test lies between 0 and 1 and is interpreted as following: 0, no association; 0.05 to 0.1, weak; 0.1 to 0.15, moderate; 0.15 to 0.25, strong; >0.25, very strong.

Results

The Fleiss kappa for all observers was almost perfect (0.815) for typical COVID-19 appearance ($P < .0001$), substantial (0.636) for indeterminate COVID-19 appearance ($P < .0001$), and almost perfect for atypical (0.806) and negative (0.962) COVID-19 appearances (see Table 2). Using Cramer V analysis, there were very strong correlations between all radiologists’ interpretations, statistically significant for all (typical, indeterminate, atypical, and negative) COVID-19 appearances ($P < .001$; see Table 3). Of the cases with overall typical imaging appearance for COVID-19, specific imaging findings

### Table 1. Frequency Distribution of Typical COVID-19 Findings.

| Typical cases out of total (303) | Rad 1 (%) | Rad 2 (%) | Rad 3 (%) |
|---------------------------------|----------|----------|----------|
| Ground glass total              | 137 (97.9%) | 157 (98.1%) | 130 (93.5%) |
| Round                           | 124 (88.6%) | 154 (96.3%) | 124 (89.2%) |
| Peripheral                      | 137 (97.9%) | 156 (97.5%) | 133 (95.7%) |
| Crazy paving total              | 55 (39.3%) | 37 (23.1%) | 44 (31.7%) |
| Round                           | 31 (22.1%) | 32 (20%) | 40 (28.8%) |
| Consolidation                   | 91 (65%) | 95 (59.4%) | 72 (51.8%) |
| Round                           | 49 (35%) | 90 (56.3%) | 61 (43.9%) |
| Peripheral                      | 80 (57.1%) | 104 (65%) | 81 (58.3%) |
| Peribronchovascular             | 45 (32.1%) | 43 (26.9%) | 48 (34.5%) |
| Perilobular                     | 70 (50%) | 49 (30.6%) | 29 (20.9%) |
| Consolidation with reverse halo | 55 (39.3%) | 20 (12.5%) | 33 (23.7%) |
| Posterior distribution          | 140 (100%) | 160 (100%) | 139 (100%) |
| Bronchial dilatation            | 5 (3.6%) | 11 (6.9%) | 6 (4.3%) |

### Table 2. Inter-Rater Agreement Between Chest Radiologists According to the RSNA Expert Consensus Statement on Reporting Chest CT Findings Related to COVID-19.

| COVID-19 appearance | Agreement | Fleiss kappa | Standard error | P value | Interpretation |
|---------------------|-----------|--------------|----------------|---------|----------------|
| Typical COVID-19 appearance |          |              |                |         |                |
| Radiologist 1 and radiologist 2 and radiologist 3 | 90.1% | 0.815 | 0.034 | <.0001 | Almost perfect agreement |
| Radiologist 1 and radiologist 2 | 91.1% | 0.821 | 0.033 | <.0001 | Almost perfect agreement |
| Radiologist 1 and radiologist 3 | 91.1% | 0.823 | 0.032 | <.0001 | Almost perfect agreement |
| Indeterminate COVID-19 appearance |          |              |                |         |                |
| Radiologist 1 and radiologist 2 and radiologist 3 | 89.1% | 0.636 | 0.063 | <.0001 | Substantial agreement |
| Radiologist 1 and radiologist 2 | 89.4% | 0.668 | 0.054 | <.0001 | Substantial agreement |
| Radiologist 1 and radiologist 3 | 89.8% | 0.641 | 0.058 | <.0001 | Substantial agreement |
| Atypical COVID-19 appearance |          |              |                |         |                |
| Radiologist 1 and radiologist 2 and radiologist 3 | 98.0% | 0.806 | 0.071 | <.0001 | Almost perfect agreement |
| Radiologist 1 and radiologist 2 | 98.3% | 0.823 | 0.074 | <.0001 | Almost perfect agreement |
| Radiologist 2 and radiologist 3 | 97.7% | 0.762 | 0.086 | <.0001 | Substantial agreement |
| Negative for pneumonia |          |              |                |         |                |
| Radiologist 1 and radiologist 2 and radiologist 3 | 98.3% | 0.962 | 0.018 | <.0001 | Almost perfect agreement |
| Radiologist 1 and radiologist 2 | 98.7% | 0.968 | 0.016 | <.0001 | Almost perfect agreement |
| Radiologist 2 and radiologist 3 | 98.3% | 0.960 | 0.018 | <.0001 | Almost perfect agreement |

Abbreviations: CT, computed tomography; RSNA, Radiological Society of North America.
including ground glass, consolidation, and crazy-paving were recorded in similar percentages of cases by the observers (see Table 1). A posterior predominance of distribution was observed in all typical cases (100%). Examples of typical COVID-19 CT appearances are presented in Figures 1 to 6.

### Discussion

COVID-19 may result in a wide spectrum of CT imaging findings, which in isolation or combination can be of variable significance. Findings may vary depending on the severity of disease, time point in the infection, or the presence of background or coexistent lung disease. As a novel infection, by definition, experience of interpreting radiologists is limited, with continuously evolving described imaging features requiring time to establish diagnostic confidence. Given the potential implications of a COVID-19 diagnosis to the patient, exposed health care workers, and the wider community, concise and unambiguous radiology reporting is of paramount importance.

The RSNA expert consensus statement on reporting chest CT findings related to COVID-19 was published to facilitate reporting of CT imaging studies acquired during the COVID-19 pandemic. The statement’s main purpose was to provide radiologists with standardized reporting language to apply to patients under investigation for COVID-19 and to allow them to assign an evidence-based impression of their level of suspicion for COVID-19 infection, thus facilitating clear communication with referring physicians. Currently, there is a paucity of data on the initial experience utilizing the RSNA expert consensus statement guidelines in clinical practice.

Our results show that there is overall good agreement for typical, indeterminate, atypical, and negative imaging findings between chest radiologists according to the RSNA expert consensus statement on reporting chest CT findings related to COVID-19. The correlation between radiologists for typical, indeterminate, atypical, and negative imaging findings is presented in Table 3.

### Table 3. Correlation Between Chest Radiologists According to the RSNA Expert Consensus Statement on Reporting Chest CT Findings Related to COVID-19.

| COVID-19 appearance       | Cramer V     | P value | Interpretation         |
|---------------------------|--------------|---------|------------------------|
| Typical COVID-19 appearance |              |         |                        |
| Radiologist 1 and radiologist 2 | 0.810       | <.001   | Very strong correlation |
| Radiologist 1 and radiologist 3 | 0.821       | <.001   | Very strong correlation |
| Radiologist 2 and radiologist 3 | 0.831       | <.001   | Very strong correlation |
| Indeterminate COVID-19 appearance |              |         |                        |
| Radiologist 1 and radiologist 2 | 0.611       | <.001   | Very strong correlation |
| Radiologist 1 and radiologist 3 | 0.669       | <.001   | Very strong correlation |
| Radiologist 2 and radiologist 3 | 0.665       | <.001   | Very strong correlation |
| Atypical COVID-19 appearance |              |         |                        |
| Radiologist 1 and radiologist 2 | 0.823       | <.001   | Very strong correlation |
| Radiologist 1 and radiologist 3 | 0.842       | <.001   | Very strong correlation |
| Radiologist 2 and radiologist 3 | 0.774       | <.001   | Very strong correlation |
| Negative for pneumonia |              |         |                        |
| Radiologist 1 and radiologist 2 | 0.960       | <.001   | Very strong correlation |
| Radiologist 1 and radiologist 3 | 0.968       | <.001   | Very strong correlation |
| Radiologist 2 and radiologist 3 | 0.960       | <.001   | Very strong correlation |

Abbreviations: CT, computed tomography; RSNA, Radiological Society of North America.

**Figure 1.** Bilateral posterior and peripheral predominant ground-glass opacities in (A) (short arrows), with a slightly more rounded appearance in (B) (short arrows).
appearances according to the RSNA expert consensus statement reporting guidelines among expert chest radiologists in suspected COVID-19 cases where RT-PCR was largely unavailable. The RSNA expert consensus statement guidelines provide both standardized language and impression categories, which will likely reduce reporting variability and improve clarity in report interpretation among referring physicians. The CO-RADS (COVID-19 Reporting and Data System) is an additional recently described categorical reporting system to assess the degree of lung involvement by COVID-19 on chest CT in patients with moderate to severe symptoms with categories 1 to 5 according to increasing suspicion on CT. The CO-RADS system was not assessed in this study but reports very good performance in predicting...
COVID-19 patients with substantial interobserver agreement on initial clinical application.18

Commonly reported imaging features with greater specificity for COVID-19 pneumonia include peripheral and bilateral ground glass and/or multifocal ground glass with rounded morphology with or without consolidation or visible intralobular lines ("crazy paving") and features of organizing pneumonia including the reverse halo sign.10,13,19 Ground-glass opacities and consolidation were recorded in a high percentage of typical COVID-19 cases among observers in our study with good agreement (see Table 1), emphasizing the importance of identifying these features on CT in patients with suspected COVID-19 where RT-PCR results may be delayed or unavailable. Posterior predominance of distribution20 has been reported in COVID-19 patients and was almost uniformly observed in the typical cases in this study. Bronchial dilatation in affected areas21 is less commonly seen in this study in line with recent literature.22 Computed tomography may be normal early on and subsequently depict an acute lung injury response to an infectious insult, with peripheral predominant multifocal ground glass with or without consolidation followed by organizing pneumonia reaction pattern.23

Due to the highly infectious nature of COVID-19 with estimated reproduction number (R0) of 2 to 2.53, rapid and
accurate diagnostic methods are needed to identify, isolate, and treat the patients in a timely fashion, which could reduce mortality rates and the risk of public contamination. In cases where RT-PCR is initially negative, typical findings on CT may encourage maintenance of infection control mechanisms and prompt repeat testing.\textsuperscript{8,24,25} In cases of high clinical suspicion but negative RT-PCR, a combination of repeated swab tests and CT imaging may be useful to reach the diagnosis and assist management.\textsuperscript{9,26} With many countries beginning to ease lockdown restrictions and health care systems returning to near pre-pandemic levels of operation, it is important that radiologists remain vigilant to the typical COVID-19 CT imaging findings with the almost inevitable anticipated second wave of the pandemic. Such typical CT findings should be kept in mind for the foreseeable future, particularly in the outpatient CT reporting setting, as it is known that asymptomatic patients with COVID-19 can have abnormalities on CT.\textsuperscript{27}

Our study has several limitations. Reverse transcription polymerase chain reaction was unavailable in the majority of cases and therefore we were only able to assess interobserver agreement according to the RSNA expert consensus statement in suspected COVID-19 cases. Furthermore, findings could not be correlated with timing of onset of infection, severity of symptoms, or patient outcome due to lack of available clinical data.

**Conclusion**

The role of chest CT imaging in the era of COVID-19 continues to evolve with a particularly important role in cases with high clinical suspicion but negative or unavailable RT-PCR. We have shown that the RSNA consensus statement on reporting chest CT findings related to COVID-19 has good interobserver agreement among expert chest radiologists in a relatively large cohort of patients with suspected COVID-19 and unavailable RT-PCR testing. The guidelines provide a framework to which radiologists can easily refer and language which can easily be applied in daily clinical practice in order to accurately communicate their level of suspicion for COVID-19 based on the presence of evidence-based objective imaging findings.

**Authors’ Note**

D. Byrne and S. B. O’Neill contributed equally to this manuscript.

**Declaration of Conflicting Interests**

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

**Funding**

The author(s) received no financial support for the research, authorship, and/or publication of this article.

**ORCID iDs**

Danielle Byrne, FFR (RCISI) https://orcid.org/0000-0002-8168-2159
C. Isabela Silva Müller, MD, PhD https://orcid.org/0000-0003-2422-2735
