Ann Med Res 2021;28(6):1217-22

Annals of Medical Research

DOI: 10.5455/annalsmedres.2020.10.1042

Original Article

COVID-19 computed tomography reporting systems: inter-rater and inter-system agreement study

Ural Koc[1], Ozlem Unal[2], Erdem Ozkan[1], Bircan Kayaaslan[3]

[1]Clinic of Radiology, Ankara City Hospital, Ankara, Turkey
[2]Department of Radiology, Ankara Yildirim Beyazit University, Ankara, Turkey
[3]Department of Infectious Diseases and Clinical Microbiology, Ankara Yildirim Beyazit University, Ankara, Turkey

Abstract

Aim: The aim of this study was to investigate agreements between COVID-19 reporting systems and radiologists.

Materials and Methods: A total of 100 laboratory-confirmed COVID-19 cases (49 males, 51 females; age range 19-88 years) were retrospectively collected. Firstly, computed tomography (CT) images were evaluated by two radiologists independently and blinded to clinical notes and laboratory and radiological reports and they gave their impressions independently according to four COVID-19 reporting systems, then all CTs were interpreted again by the two radiologists for extracting CT features at the same session by consensus.

Results: Bilateral, lower lobe, peripheral, dorsal and multifocal lung involvements were predominantly seen, and ground-glass opacities (GGOs) were the most common CT imaging finding in the current study. Reporting systems showed fair to moderate agreements between senior and junior raters (0.246-0.490, p<0.001). According to the assigned three-category coding system as similar to that of the Radiological Society of North America (RSNA) Expert Consensus Statement on Reporting and other reporting systems and guidance, strength of inter-rater agreement values was increased (0.365-0.576, p<0.001) and inter-system agreements were substantial to almost perfect in both raters.

Conclusion: Radiology reporting including frequently seen CT features and lung parenchyma distributions with systems based on fewer categories may provide good agreement between observers in patients with suspected COVID-19.

Keywords: Coronavirus; COVID-19; radiology information systems; tomography

INTRODUCTION

Coronavirus 2019 disease (COVID-19), which officially started in Wuhan, China, was declared a pandemic by the World Health Organization (WHO) on 11 March 2020 (1,2). Laboratory and imaging methods are important diagnostic tools in COVID-19. Reverse transcription-polymerase chain reaction (RT-PCR) is the gold standard diagnostic method for COVID-19, while computed tomography (CT) has been widely used throughout the world as a supportive imaging tool in diagnosis, management and follow-up during the course of COVID-19. The reasons for this situation are high sensitivity of CT imaging, variable sensitivity and long turnaround time for RT-PCR test results (3). Although it is not routinely recommended (4-5), we observed that it is used as a screening tool in many countries with high prevalence of COVID-19.

After a certain amount of imaging data accumulation in their single centers, researchers, especially in China, have revealed typical and atypical CT imaging findings (6). Various radiological reporting systems and guidelines such as the Radiological Society of North America (RSNA) Expert Consensus Statement on Reporting, COVID-19 Reporting and Data System (CO-RADS), COVID-19 Imaging Reporting and Data System (COVID-RADS), and British Society of Thoracic Imaging (BSTI) Reporting Guidance including templates, lexicons and languages have emerged to provide a structural report to decrease variability between radiologists and effective communication with clinicians in order to guide them correctly, especially in cases with clinical suspicion of COVID-19 (7-11). These systems, guidelines and templates categorize whether the cases are not exact but partial likelihood for COVID-19 according to CT image characteristics (7-10). On this point of radiological structure reporting approaches to COVID-19, we hypothesized if there were an agreement between reporting systems and radiologists, and furthermore, if the systems were simpler, as similar to the

Received: 10.10.2020 **Accepted:** 02.12.2020 **Available online:** 24.06.2021
Corresponding Author: Ural Koc, Clinic of Radiology, Ankara City Hospital, Ankara, Turkey
E-mail: dr_uralkoc@hotmail.com

RSNA Expert Consensus Statement on Reporting, then agreements could be increased.

In this study, the aim was to evaluate agreements between the radiology reporting systems and inter-raters with original recommendations and the assigned three-category coding system as in the RSNA Expert Consensus Statement.

MATERIALS and METHODS

We obtained approval from the Turkish Ministry of Health as a requirement on 21 May 2020. We received the local ethics committee approval from Ankara City Hospital Ethical Committee No.1 (approval number: E1/664/2020) on 28 May 2020. Due to the retrospective nature of this study, informed consent was waived.

Cases

Cases with laboratory-confirmed COVID-19, who were admitted to the infectious disease clinic, were retrospectively screened by O.U. and B.K. 100 cases were collected. Evaluation was made of 100 cases (49 males, 51 females) with a median age of 42 years (range, 19-88 years) by two radiologists. Nasopharyngeal swab testing and CT acquisitions were on the same day. All nasopharyngeal swab testing results were positive for COVID-19.

Imaging Technique

Cases underwent a chest CT examination with a 128 slice Revolution Evo CT scanner (GE Healthcare). The body region between lung apices to the inferior level of the costophrenic angle was included at full inspiration in supine position. Intravenous (IV) contrast material was not used for CT acquisition. CT acquisition parameters were slice thickness: 1.25 mm; tube voltage: 100-120 kV; 80-400 mAs; 1.375 pitch; 0.625 reconstruction interval; and 0.5 seconds (sec) rotation time. The tube current was regulated with automatic exposure control (AEC) system (ASIR-V, GE, Healthcare).

Imaging Interpretation

Two radiologists (U.K.-senior- and E.O.-junior-, with 8 and 2 years of experience with chest CT, respectively) evaluated the cases independently and blinded to clinical notes and laboratory and radiological reports in the picture archiving and communication systems (PACS). The lung window settings were used (window width: 1500 hounsfield units (HU), window level: -600 HU). Multiplanar reconstruction (MPR) techniques were also used in the image interpretation sessions. The senior radiologist tried to follow all recent radiology literature and gain experience related to COVID-19. The junior radiologist was able to follow the literature of COVID-19 radiology partially. The two radiologists interpreted the chest CT by four reporting systems blinded to each other (7-10). All reporting systems were assigned with a three-category coding system in order to provide a standardization which was similar to that of the RSNA Expert Consensus Statement on Reporting (Table 1). After independent interpretation for categorization, the chest CT findings including lung lobe involvements, lesion characteristics and distributions (anterior-posterior, transverse, scattered, side) were reviewed.

Table 1. Assigned code attributes were shown

CODE	CO-RADS (7)	RSNA Consensus Statement on Reporting (8)	COVID-RADS (9)	BSTI Reporting Guidance (10)
A	2	Atypical	1	Non-Covid
I	3	Indeterminate	2A	Indeterminate
T	4 and 5	Typical	2B and 3	Classical and Probable

CO-RADS, COVID-19 Reporting and Data System; COVID-RADS, COVID-19 Imaging Reporting and Data System; RSNA, Radiological Society of North America; BSTI, British Society of Thoracic Imaging

Statistical Analysis

Since the number of cases is a hundred, the numbers indicated also reflect the frequency. The age parameter was given as median (minimum-maximum). Agreements between the senior and junior radiologist, and reporting systems were evaluated by Cohen's kappa statistics. Spearman Correlation test was applied for correlation analysis between reporting systems. We used the SPPS for Windows software package (version 21.0, SPSS Inc.) for statistical analysis and accepted p value <0.05 as statistically significant.

RESULTS

Demographics and CT Features

A total of 100 cases were included in the study and median age was 42 years with a range of 19-88, of which 51 were female. According to the imaging findings, bilateral, lower lobe, peripheral, dorsal and multifocal lung involvements were predominantly seen in the current study. Ground-glass opacities (GGOs) were the most common CT feature amongst others. All CT features are given in Table 2 and classified according to the assigned codes in Table 3.

Table 2. Basic demographics and CT features of cases in the study

Findings	All Cases (n=100)
Age(y)	42(19-88)
Sex	
Male	49
Female	51
Imaging Findings	
Frequency of lobe involvement	
Right upper lobe	45
Left upper lobe	43
Middle lobe	41
Right lower lobe	70
Left lower lobe	58
Lung region distribution	
Bilateral	56
Unilateral	44

Transverse distribution	
Central	4
Peripheral	58
Peripheral and Central	38
Anterior-posterior distribution	
Ventral	12
Dorsal	44
Ventral and Dorsal	44
Scattered distribution	
Focal/Single	38
Multifocal	62
Ground-glass opacities(GGO)	90
Consolidation	28
GGO with superimposed consolidation	25
Crazy paving patern	15
Reversed halo sign	11
Vascular enlargement	32
Subpleural curvilinear bands	7
Centrilobular nodular pattern	10
Air bronchogram	11
Bronchiectasis	9
Bronchial wall thickening	6
Halo sign	14
Pleural effusion	4
Pleural thickening	3
Lymphadenopathy	6
Pericardial effusion	4
Pulmonary emphysema	11

GGOs, Ground-glass opacities

Table 3. The chest CT features were classified according to the assigned codes

A	I	T
Pleural effusion	Unilaterality	Bilaterality
Pleural thickening	Upper or Middle lobe predominance	Lower lobe predominance
Lymphadenopathy	Central distribution	Peripheral distribution or peripheral predominance
Pericardial effusion	Ventral distribution	Dorsal distribution or dorsal predominance
Pulmonary emphysema	Focal/Single lesion	Multifocality
Bronchiectasis	Consolidation	Ground-glass opacities(GGO)
Centrilobular nodular pattern	Air bronchogram	GGO with superimposed consolidation
Halo sign	Bronchial wall thickening	Crazy paving pattern
	Vascular enlargement	Reversed halo sign
	Subpleural curvilinear bands	

Inter-rater Agreement and Discrepancy

Reporting systems showed fair-moderate agreements between raters (0.246-0.490, p<0.001) (Table 4). The highest discrepancy was viewed in COVID-RADS and CO-RADS systems. According to the three-category coding system described in Table 1, inter-rater agreements were moderate (0.365-0.576, p<0.001) (Table 5), but showed better strength of agreement values than the previous ones given in Table 4. Discrepancies decreased slightly in COVID-RADS, CO-RADS and BSTI reporting guidance systems (Table 5).

Table 4. Inter-rater agreement with different radiological reporting systems, statement and guidance

	Senior Rater	Junior Rater	Cohen's Kappa (κ)	p value
CO-RADS				
2	10	13		
3	28	42		
4	8	25	0.380	<0.001
5	54	20		
COVID-RADS				
1	8	16		
2A	18	36	0.246	<0.001
2B	8	26		
3	66	22		
RSNA Consensus Statement on Reporting				
Atypical	10	20		
Indeterminate	38	45	0.490	<0.001
Typical	52	35		
BSTI Reporting Guidance				
Non-COVID	10	12		
Indeterminate	48	42		
Probable	4	25	0.426	<0.001
Classic	48	21		

CO-RADS, COVID-19 Reporting and Data System; COVID-RADS, COVID-19 Imaging Reporting and Data System; RSNA, Radiological Society of North America; BSTI, British Society of Thoracic Imaging

Table 5. Inter-rater agreement with assigned code to different radiological reporting systems

	CO-RADS			
Assigned Code	Senior Rater	Junior Rater	Cohen's Kappa (κ)	p value
A	10	13		
I	28	42	0.543	<0.001
T	62	45		
	COVID-RADS			
Assigned Code	Senior Rater	Junior Rater	Cohen's Kappa (κ)	
A	8	16		
I	18	36	0.365	<0.001
T	74	48		
	RSNA Consensus Statement on Reporting			
Assigned Code	Senior Rater	Junior Rater	Cohen's Kappa (κ)	
A	10	20		
I	38	45	0.49	<0.001
T	52	35		
	BSTI Reporting Guidance			
Assigned Code	Senior Rater	Junior Rater	Cohen's Kappa (κ)	
A	10	12		
I	38	42	0.576	<0.001
T	52	46		

CO-RADS, COVID-19 Reporting and Data System; COVID-RADS, COVID-19 Imaging Reporting and Data System; RSNA, Radiological Society of North America; BSTI, British Society of Thoracic Imaging

Table 6. Correlations and inter-system agreements

	Senior Rater		Junior Rater	
	RSNA Consensus Statement on Reporting		RSNA Consensus Statement on Reporting	
	r	p	r	p
CO-RADS	0.935	<0.001	0.866	<0.001
COVID-RADS	0.795	<0.001	0.857	<0.001
BSTI Reporting Guidance	0.982	<0.001	0.874	<0.001
	Cohen's Kappa (κ)		Cohen's Kappa (κ)	
CO-RADS	0.822	<0.001	0.745	<0.001
COVID-RADS	0.573	<0.001	0.639	<0.001
BSTI Reporting Guidance	1	<0.001	0.696	<0.001

CO-RADS, COVID-19 Reporting and Data System; COVID-RADS, COVID-19 Imaging Reporting and Data System; RSNA, Radiological Society of North America; BSTI, British Society of Thoracic Imaging

Reporting Systems Agreement

There was a positive correlation between systems with original definitions and groupings both in the senior (0.795-0.935, p<0.001) and junior rater (0.857-0.874, p<0.001) (Table 6). According to the three-category coding system described in Table 1, inter-system agreements were almost perfect between the RSNA Consensus Statement on Reporting, CO-RADS and BSTI Reporting Guidance (0.822-1, p<0.001) in the senior rater, while there were substantial agreements (0.639-745, p<0.001) in the junior rater (Table 6).

DISCUSSION

Radiology reports, which contain clear impressions, common terminology and lexicon rather than describing CT findings, are more beneficial for requesting physicians and also contribute to the efficiency of the radiologist in making an evaluation in patient management (11-13). Good quality radiological data collected by creating this basis can be used in the area of epidemiology, research and development (11-14). A recently published systematic review and meta-analysis including 1115 patients showed that CT features may be helpful in disease course and management of COVID-19 patients (15). Many leading specialist and subspecialist radiological societies and associations have declared their consensus statements, algorithms, recommendations and guidelines related to imaging and reporting in the era of COVID-19 (5,8,16-19). Moreover, the "COVID-19 Standardized Reporting" working group of the Dutch Radiological Society introduced

CO-RADS (7), while Salehi et al. developed COVID-RADS based on a large group of imaging studies (9). In the current study, it was investigated how the most appropriate reporting impression should be during this period that offers so many partially similar reporting proposals, which are sometimes confusing for radiologists. To best of our knowledge, there has not yet been any study related to agreements between different reporting proposals.

Large sample-size systematic reviews and meta-analyses regarding the CT finding distributions of COVID-19 revealed that typical spatial parameters were predominantly lower lobe involvement, especially right lower lobe, bilaterality, multifocality, and posterior and peripheral distribution (6,15,20). On the other hand, GGOs and mixed patterns of GGOs and consolidations were frequently seen CT features (6,15). Knowing those spatial distributions and common CT features provides great sensitivity for diagnostic test accuracy, evaluation and management of COVID-19 (6,15,20). In this current study, the predominant CT manifestations included right lower lobe involvement (70%), bilaterality (56%), and peripheral (58%), dorsal (44%), dorso-ventral (44%) and multifocal (62%) distributions, which are compatible with the literature. Other typical, indeterminate and atypical findings were defined and reported at various frequencies in the literature (21), though these were observed less frequently in patients with laboratory confirmed COVID-19 than in the dominant findings we mentioned earlier.

Structured recommendation report formats and systems have been published and shared by many societies in order to effectively utilize the detected radiological findings to be useful in the clinical field and to increase productivity of radiologists (7-10). The RSNA Expert Consensus Statement on Reporting and BSTI Reporting Guidance show substantially similar categorization with different category names, but there is an intermediate layer named as probable COVID-19 between classic and indeterminate groups in BSTI Reporting Guidance (8,10). CO-RADS includes six categories which are introduced with definitions, spatial distribution parameters and CT imaging features (7). COVID-RADS provides grade and level of suspicion in light of CT findings except distribution pattern (only focality or multifocality are included) which is defined as atypical, fairly typical, a combination of atypical findings with typical findings, and typical findings (9). We found positive high correlation between the RSNA Expert Consensus Statement on Reporting and the other proposals, especially with CO-RADS and BSTI Reporting Guidance in both raters.

The CO-RADS study revealed that the authors observed 28% discrepancy in a single CO-RADS category; moreover, kappa values showed moderate-substantial agreement between each observer and median of the other observers (7). Our results showed that the RSNA Expert Consensus Statement on Reporting had higher values of kappa in contrast to the others between senior and junior raters. After assigning codes for pairing categories into a single category, which was similar to the RSNA Expert Consensus Statement on Reporting, kappa values increased in CO-RADS, COVID-RADS and BSTI Reporting Guidance. Our results suggested that concise reporting systems decrease variability and discrepancy. COVID-RADS has been the most recently introduced one among others and inter-rater and inter-system agreement were lower than the others; however, high level of suspicion (typical findings) cases were 66% in COVID-RADS, which was higher than the other systems in the senior rater assessment.

Furthermore, the assigned coding system to reduce categories provided almost perfect-substantial agreement between the RSNA Expert Consensus Statement on Reporting, CO-RADS and BSTI Reporting Guidance, and an increase in detected cases with typical findings.

LIMITATIONS

The study has some limitations. The senior and junior authors were blinded to clinical and laboratory data, which highly impact image interpretations in routine. Intra-rater experience in these different reporting systems and inter-system effect may influence the rater impressions related to categories.

CONCLUSION

Reporting systems with defined frequently seen comprehensive CT features and lung distributions, and systems with fewer categories may provide strong agreements between observers in patients with suspected COVID-19.

Competing interests: The authors declare that they have no competing interest.

Financial Disclosure: There are no financial supports.

Ethical approval: We received the local ethics committee approval from Ankara City Hospital Ethical Committee No.1, Turkey (approval number: E1/664/2020) on 28 May 2020.

REFERENCES

1. World Health Organization. Coronavirus disease 2019 (COVID-19) situation report-1. World Health Organization, Geneva. Available via https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200211-sitrep-22-ncov.pdf?sfvrsn=fb6d49b1_2. Accessed 11 Jun 2020
2. World Health Organization. Coronavirus disease 2019 (COVID-19) situation report-51. World Health Organization, Geneva. Available via https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200311-sitrep-51-covid-19.pdf?sfvrsn=1ba62e57_10. Accessed 11 Jun 2020
3. Ai T, Yang Z, Hou H, et al. Correlation of chest CT and RT-PCR testing in coronavirus disease 2019 (COVID-19) in China: a report of 1014 cases. Radiology 2020;296:32-40.

4. American College of Radiology (ACR). ACR recommendations for the use of chest radiography and computed tomography (CT) for suspected COVID-19 infection. Available via https://www.acr.org/Advocacy-and-Economics/ACR-Position-Statements/Recommendations-for-Chest-Radiography-and-CT-for-Suspected-COVID19-Infection. Accessed 11 Jun 2020.
5. Rubin GD, Ryerson CJ, Haramati LB, et al. The role of chest imaging in patient management during the COVID-19 pandemic: a multinational consensus statement from the Fleischner Society. Radiology 2020;296:172-80.
6. Ojha V, Mani A, Pandey NN, et al. CT in coronavirus disease 2019 (COVID-19): a systematic review of chest CT findings in 4410 adult patients. Eur Radiol 2020.
7. Prokop M, van Everdingen W, van Rees Vellinga T. CO-RADS – a categorical CT assessment scheme for patients with suspected COVID-19: definition and evaluation. Radiology 2020;30:4930-42.
8. Simpson S, Kay FU, Abbara S, et al. Radiological Society of North America expert consensus statement on reporting chest CT findings related to COVID-19. Endorsed by the Society of Thoracic Radiology, the American College of Radiology, and RSNA. J Thorac Imaging 2020;35:219-27.
9. Salehi S, Abedi A, Balakrishnan S, et al. Coronavirus disease 2019 (COVID-19) imaging reporting and data system (COVIDRADS) and common lexicon: a proposal based on the imaging data of 37 studies. Eur Radiol 2020.
10. The British Society of Thoracic Imaging. Covid-19 Resources. Available via https://www.bsti.org.uk/covid-19-resources/covid-19-bsti-reporting-templates. Accessed 11 Jun 2020
11. Schwartz LH, Panicek DM, Berk AR, et al. Improving communication of diagnostic radiology findings through structured reporting. Radiology 2011;260:174-81.
12. Noumeir R. Benefits of the DICOM structured report. J Digit Imaging 2006;19:295-306.
13. Ganeshan D, Duong PT, Probyn L, et al. Structured Reporting in Radiology. Acad Radiol 2018;25:66-73.
14. Hoore J, Barnett I, Boland MR, et al. Ideas for how informaticians can get involved with COVID-19 research. BioData Min 2020;13:3.
15. Wan S, Li M, Ye Z, et al. CT manifestations and clinical characteristics of 1115 patients with coronavirus disease (COVID-19): A systematic review and meta-analyis. Acad Radiol 2020;27:910-21.
16. Vogel-Clausen 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 2020;192:633-40.
17. Nair A, Rodrigues JCL, Hare S, et al. A British Society of Thoracic Imaging statement: considerations in designing local imaging diagnostic algorithms for the COVID-19 pandemic. Clin Radiol 2020;75:329-34.
18. Dennie C, Hague C, Lim RS, et al. Canadian Society of Thoracic Radiology/Canadian Association of Radiologists consensus statement regarding chest imaging in suspected and confirmed COVID-19. Can Assoc Radiol J 2020;71:470-81.
19. Revel MP, Parkar AP, Prosch H, et al. COVID-19 patients and the radiology department-advice from the European Society of Radiology (ESR) and the European Society of Thoracic Imaging (ESTI). Eur Radiol 2020;30:4903-9.
20. Xu B, Xing Y, Peng J, et al. Chest CT for detecting COVID-19:a systematic review and meta-analyis of diagnostic accuracy. Eur Radiol 2020;81:16-25.
21. Ufuk F, Savas R. Chest CT features of the novel coronavirus disease (COVID-19). Turk J Med Sci 2020;50:664-78.