Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Clinical Imaging 94 (2023) 1–8

Contents lists available at ScienceDirect

Clinical Imaging

journal homepage: www.elsevier.com/locate/clinimag

Cardiothoracic Imaging

Can a chest HRCT-based crash course on COVID-19 cases make inexperienced thoracic radiologists readily available to face the next pandemic?☆

Lorenzo Cereser[a,*], Emanuele Passarotti[a], Annarita Tullio[b], Vincenzo Patruno[c], Leonardo Monterubbiano[a], Pierpaolo Apa[a], Chiara Zuiani[a], Rossano Girometti[a]

[a] Institute of Radiology, Department of Medicine, University of Udine, University Hospital "S. Maria della Misericordia", p.le S. Maria della Misericordia, 15, 33100 Udine, Italy
[b] Institute of Hygiene and Clinical Epidemiology, University Hospital "S. Maria della Misericordia", p.le S. Maria della Misericordia, 15, 33100 Udine, Italy
[c] Pulmonology Department, "S. Maria della Misericordia" University Hospital, p.le S. Maria della Misericordia, 15, 33100 Udine, Italy

ARTICLE INFO

Keywords:
COVID-19
Pandemics
Radiologists
Education

ABSTRACT

Objective: To test the inter-reader agreement in assessing lung disease extent, HRCT signs, and Radiological Society of North America (RSNA) categorization between a chest-devoted radiologist (CR) and two HRCT-naïve radiology residents (RR1 and RR2) after the latter attended a COVID-19-based chest high-resolution computed tomography (HRCT) "crash course".

Methods: The course was built by retrospective inclusion of 150 patients who underwent HRCT for COVID-19 pneumonia between November 2020 and January 2021. During a first 10-days-long "training phase", RR1 and RR2 read a pool of 100/150 HRCTs, receiving day-by-day access to CR reports as feedback. In the subsequent 2-days-long "test phase", they were asked to report 50/150 HRCTs with no feedback. Test phase reports of RR1/RR2 were then compared with CR using unweighted or linearly-weighted Cohen's kappa (k) statistic and intraclass correlation coefficient (ICC).

Results: We observed almost perfect agreement in assessing disease extent between RR1-CR ($k = 0.83$, $p < 0.001$) and RR2-CR ($k = 0.88$, $p < 0.001$). The agreement between RR1-CR and RR2-CR on consolidation, crazy paving pattern, organizing pneumonia (OP) pattern, and pulmonary artery (PA) diameter was substantial ($k = 0.65$ and $k = 0.68$), moderate ($k = 0.42$ and $k = 0.51$), slight ($k = 0.10$ and $k = 0.20$), and good-to-excellent (ICC = 0.87 and ICC = 0.91), respectively. The agreement in providing RSNA categorization was moderate for R1 versus CR ($k = 0.56$) and substantial for R2 versus CR ($k = 0.67$).

Conclusion: HRCT-naïve readers showed an acceptable overall agreement with CR, supporting the hypothesis that a crash course can be a tool to readily make non-subspecialty radiologists available to cooperate in reading high burden of HRCT examinations during a pandemic/epidemic.

1. Introduction

Depending on variant dominance, geographic location, and underlying chronic disease, the Coronavirus Disease 2019 (COVID-19) presents a broad spectrum of clinical manifestations, ranging from mild upper airway symptoms to severe pneumonia requiring mechanical ventilatory support.[1,2]

Chest high-resolution computed tomography (HRCT) has gained a role in supporting clinical decision-making in COVID-19 pneumonia, e.g., for assessing pulmonary complications, follow-up, and prognostic evaluation.[3] Previous studies demonstrated that HRCT features at admission, including lung disease extent, consolidation, crazy-paving pattern, organizing pneumonia (OP) pattern, and pulmonary artery (PA) diameter, are associated with adverse outcomes, such as hospitalization, intensive care unit admission, and death.[4–10]

Not surprisingly, chest HRCT is a highly requested examination in

☆ This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.
* Corresponding author.
E-mail addresses: lcereser@sirm.org (L. Cereser), vincenzo.patruno@asufc.sanita.fvg.it (V. Patruno), chiara.zuiani@uniud.it (C. Zuiani), rgirometti@sirm.org (R. Girometti).

https://doi.org/10.1016/j.clinimag.2022.11.010
Received 8 August 2022; Received in revised form 1 November 2022; Accepted 14 November 2022
Available online 22 November 2022

the current scenario and has been even more frequently requested at the beginning of the pandemic and during subsequent waves. Within this framework, many non-thoracic radiologists have been called to report a high burden of HRCT examinations while having limited or no subspecialty skills and, in turn, examining unusual, unprecedented HRCT findings. One might assume that a similar scenario might repeat if the pandemic rises again or new future pandemics occur. This prompts the problem of how to make non-thoracic radiologists able to evaluate HRCT examinations in an "emergency" setting accurately and in line with recommendations on how to report and categorize imaging findings, e.g., those by the Radiological Society of North America (RSNA) in the case of COVID-19 pneumonia.[11,12]

Based on previous proof of educational success in different radiology fields,[13,14] we hypothesized that attending an intensive training program (so-called "crash course") might provide non-skilled radiologists with the minimum expertise required to recognize and quantify lung involvement in a frontline scenario. We developed a crash course model for HRCT-naïve radiologists based on COVID-19 HRCT examinations, including educational material and training case presentation of 100 HRCTs encompassing the whole spectrum of lung disease severity, and then a test on a different pool of 50 cases. We hypothesized that, if proven effective, this strategy might be used even in future pandemics from different agents. As far as we know, no previous similar studies have been done in this setting.

This study aimed to test a crash course model by evaluating the inter-reader agreement in assessing HRCT lung disease extent between a chest-devoted radiologist (CR) and two readers with no specific experience in thoracic imaging after they attended the crash course. Secondarily, we aimed to determine the inter-reader agreement between CR and readers in assessing HRCT signs associated with unfavorable

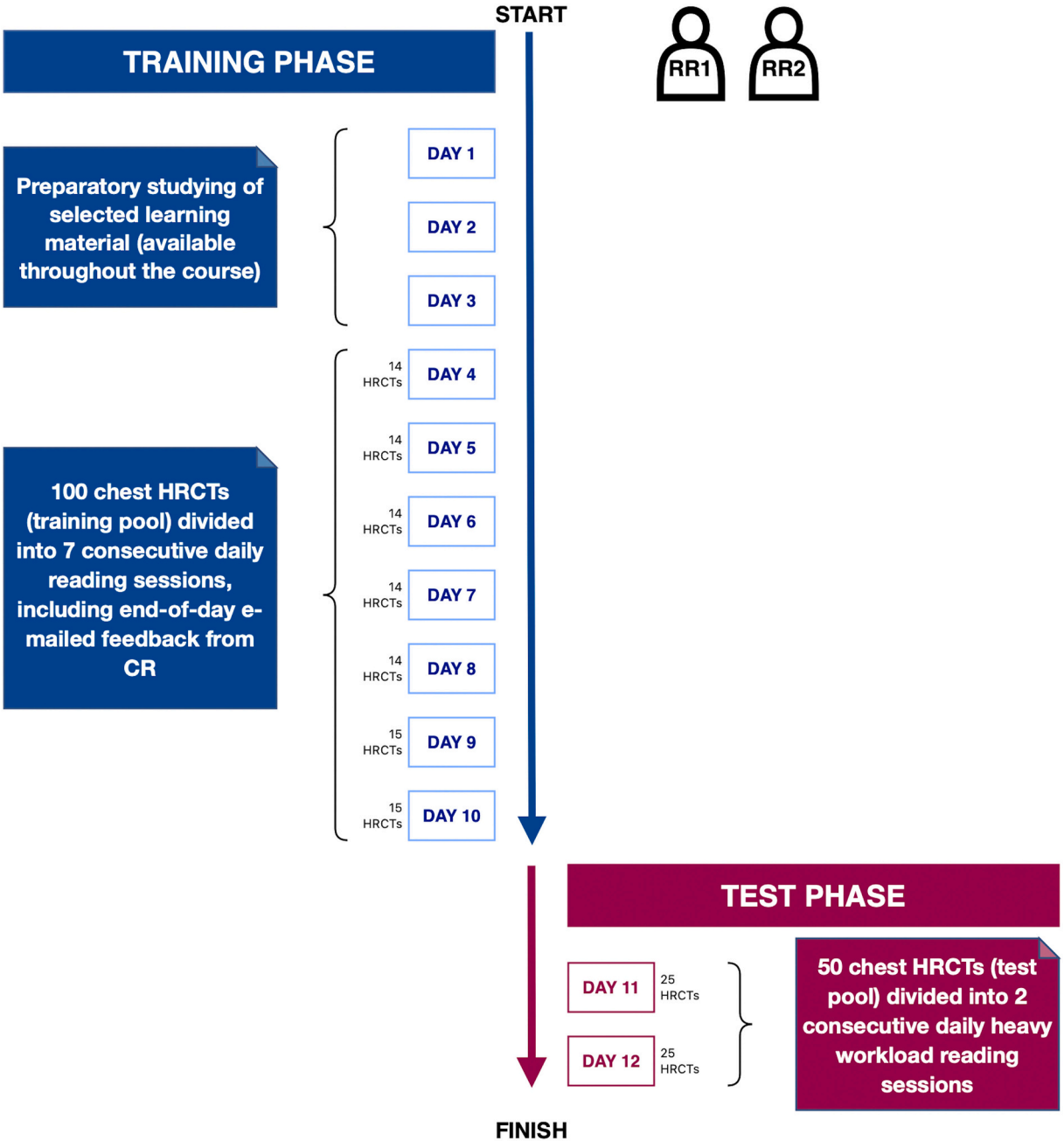


Fig. 1. Crash course design and schedule.
Footnotes: *RR1*, radiology resident 1; *RR2*, radiology resident 2; *CR*, chest-devoted radiologist.

prognosis and RSNA categorization. Overall, we hypothesized that high inter-reader agreement could prove that the educational model is effective.

2. Material and methods

2.1. Study design

Our Institutional Review Board approved this retrospective study and waived the acquisition of the informed consent from patients.

Overall, the study was coordinated by a radiologist not involved in image reading and designed as follows (Fig. 1): two radiology residents at the beginning of their first year of the residency program at our University Hospital (radiology resident 1 [RR1] and radiology resident 2 [RR2]), with no previous experience in computed tomography imaging, underwent a 12-days long crash-course based on HRCTs of COVID-19 pneumonia, which was articulated into two phases. In the first 10-days-long phase, RR1 and RR2 read a "training pool" of HRCT images, receiving day-by-day feedback on readings ("training phase" of the study). Subsequently, they were asked to report a "test pool" of HRCT cases simulating clinical activity, with no feedback on readings (2-days long "test phase" of the study). We then compared RR1 and RR2 reports of the test phase with readings performed by a chest radiologist (CR) with 13 years of experience in the field, who evaluated the training pool and test pool cases in advance. Details are explained below.

2.2. Study population, standard of reference, and HRCT pools

As mentioned, the standard of reference was represented by HRCT readings performed by CR, blinded to clinical data, who was asked to categorize HRCT examinations as negative, mild, moderate, or severe according to the computed tomography involvement score (CTIS) (Fig. 2),[15,16] and report using a COVID-19-dedicated structured model modified from Neri et al.[12] (Supplementary material 1).

We planned that the set of examinations to be read by CR should be composed of 150 HRCTs, and in turn, should include two different pools on a 2:1 rate, namely the training pool (100 cases) and test pool (50 cases) to be later presented to RR1 and RR2. The training pool was planned to include 8/100 negative, 22/100 mild, 35/100 moderate, and 35/100 severe pneumonia cases. The test pool was planned to include 50 HRCTs with case repartition mirroring the prevalence of disease severity in clinical practice,[17,18] i.e., 2 (4%) negative cases, 20 (40%) mild cases, 20 (40%) moderate cases, and 8 (16%) severe cases.

To build the set and allocate the proper number of cases to each pool based on the severity classification, we searched our COVID-19 center's database to find all the consecutive adult patients with suspected COVID-19 pneumonia who underwent HRCT and reverse transcriptase polymerase chain reaction (RT-PCR) test for SARS-CoV-2 in nasal-pharyngeal swab between November 2020 and January 2021. Patients were selected consecutively on a stepwise basis, excluding those showing negative RT-PCR tests ($n = 8$) or incomplete clinical data after HRCT ($n = 4$). Examinations meeting inclusion criteria were read by CR until the pre-planned number of cases of each severity classification slot for each pool was met (e.g., any further consecutive "negative" case was no longer read and included after the "negative examination" slot of a certain pool was filled). A total of 220 readings was needed to reach the required sample size of 150 patients. For patients who underwent multiple HRCTs, we selected the baseline one.

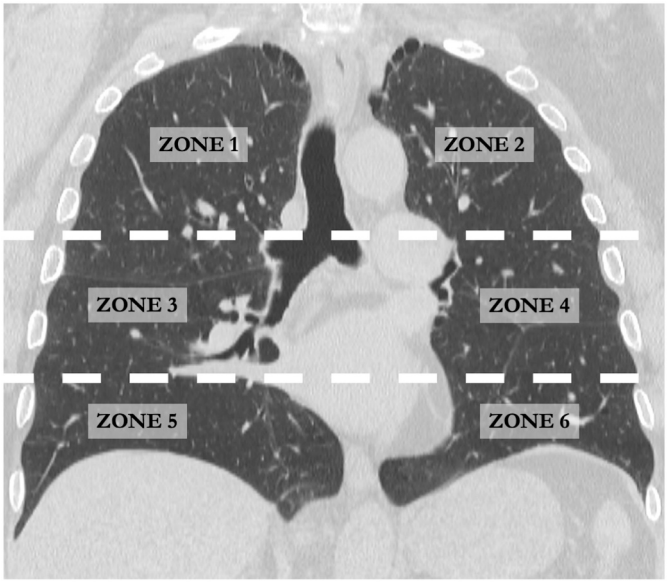
2.3. HRCT examination technique

HRCTs were performed on a 64-row Computed Tomography scanner (SOMATOMgo.Top, Siemens Healthineers, Erlangen, Germany), with the patient in supine position and during inspiratory breath-hold. The volumetric acquisition was executed as follows: tube voltage, 120 kV; tube current modulation range, 100–350 mA; reconstructed section thickness and interval, 1.50 mm. The reconstruction set included images with pulmonary parenchyma windowing and high-spatial-frequency algorithm (level, −500 Hounsfield Units [HU]; width, 1700 HU) and images with soft tissue windowing (level, 50 HU; width, 350 HU) and algorithm.

2.4. Study phases and image analysis

After the training pool and test pool were built, the order of case presentation in each pool was randomized using freely available software (https://www.randomizer.org).

The training phase then started on day 1 with a 30 min lesson by CR



SIX LUNG ZONES	LOCATION (right and left lung)
1 and 2	Above the carina
3 and 4	Between the carina and the inferior pulmonary veins
5 and 6	Below the inferior pulmonary veins

ZONAL SCORE	ZONAL INVOLVEMENT
0	0%
1	<25%
2	25-49%
3	50-75%
4	>75%

TOTAL CTIS (zonal score x 6)	LUNG DISEASE EXTENT
0	Negative
1-5	Mild
6-11	Moderate
11-24	Severe

Fig. 2. Scheme to assess the lung disease extent at HRCT according to the computed tomography involvement score (CTIS). The disease extent is estimated by dividing the lungs into six lung zones, scoring the lung involvement for each zone, and summing up all the zone scores. According to CTIS, HRCTs are classifiable as negative, or with mild, moderate, or severe lung disease extent.
Footnotes: References (15) Xie X et al., Radiology 2020; (16) Ooi GC et al., Radiology 2004.

illustrating the crash course scheduling and objectives, as well as CTIS, RSNA categorization criteria, and the structured reporting template to be used. RR1 and RR2 were subsequently provided with educational material to be consulted on days 1–3. The material consisted of selected referral papers[3,11,19–21] covering the basics of chest HRCT and COVID-19 imaging, as well as a scheme illustrating how to quantify disease extent with the CTIS (Fig. 2). This material was made available for consultation throughout the training and test phases. On days from 4 to 10, RR1 and RR2 independently evaluated training pool HRCTs during daily reading sessions, including 14 cases each on days 4 to 8, and 15 cases each on days 9 and 10. The study coordinator emailed the corresponding CR reports to RR1 and RR2 as feedback at the end of daily readings. Daily sessions (reading and feedback evaluation) were scheduled to correspond to a regular 8-h working day.

From day 11 to 12, we ran the subsequent test phase, in which RR1 and RR2 independently evaluated test pool HRCTs. To approximate a pandemic wave scenario, readers were prompted with 25 examinations per day. RR1 and RR2 did not receive CR reports during this phase.

In both study phases, RR1 and RR2 evaluated all the HRCTs on a dedicated workstation (Suitestensa Ebit srl, Esaote Group Company, Genoa, Italy), and were asked to report them with the same structured reporting template used by CR. The field "clinical indication" was prefilled with the standardized wording "suspected COVID-19 pneumonia" without additional data.

2.5. Statistical analysis

We reported study population characteristics using descriptive statistics, assessing the normality of the distributions with the Kolmogorov–Smirnov test.

For disease extent estimation, we investigated the inter-reader agreement between RR1-CR and RR2-CR in attributing the CTIS on test pool examinations. Since the disease extent was considered as a three-level categorical variable, i.e., negative/mild (CTIS <6), moderate (CTIS 6–11), and severe (CTIS >11), we used the linearly weighted Cohen's kappa (k) as per the Cicchetti-Allison method.[22] The 95% confidence interval (95%CI) has been calculated as bias-corrected and accelerated bootstrap confidence interval based on 1000 bootstrap replicates.

The secondary outcome was the evaluation of RR1-CR and RR2-CR agreement in assessing HRCT signs associated with unfavorable prognosis. For the presence of consolidation, crazy-paving pattern, and OP pattern, we used the unweighted Cohen's k for dichotomous rating variables between two raters. For the assessment of PA diameter, which is a continuous variable, we used the intraclass correlation coefficients (ICC), employing a "two-way, random effects, absolute agreement, single rater/measurement" ICC model.[23] For the RSNA categorization we used the linearly weighted Cohen's kappa (k).[22]

To test if the disease severity influenced the inter-reader agreement, we calculated the agreement between RR1-CR and RR2-CR on all the HRCT variables, i.e., disease extent, HRCT signs, and RSNA categorization, according to disease extent (dichotomized as negative-to-mild versus moderate-to-severe). The interpretation of k coefficient and ICC values was according to Landis and Koch[24] and Koo and Li,[25] respectively, as detailed in Supplementary material 2. When paradox k was observed (i.e., percent agreement >0.80, and k < 0.40) and both Prevalence Index and Bias Index were different from zero, the imbalance was corrected by using the prevalence-adjusted bias-adjusted kappa (PABAK) statistic.

As a proof of clinical utility, we used binomial regression models to test the ability of all the HRCT signs found by CR (lung disease extent, consolidation, crazy-paving pattern, OP pattern, and PA diameter) to predict "hospitalization" (hospitalization versus no hospital admission) and "intensive care unit (ICU) admission" (ICU admission versus no hospital admission or non-ICU hospitalization). Clinical utility analysis was performed in the whole set of 150 patients' readings by CR.

Statistical analyses were performed using R software 3.4.2 (R Foundation for Statistical Computing, Vienna, Austria). The significance level was 0.05.

3. Results

3.1. Patients and HRCT findings

Patient-related clinical characteristics and outcomes are presented in Table 1. Most patients (105/150, 70%) had one or more comorbidities at hospital admission, mainly cardiovascular (82/150, 55%). Nearly 70% (107/150) of patients were hospitalized, including 29 subjects admitted to ICU. Seven out of 150 patients (5%) died after being admitted to ICU.

Table 2 shows how readers assessed disease extent, HRCT signs, and RSNA categorization in the test pool. RR1 and RR2 quantified disease extent comparably to CR, with overlapping mean and median scores. According to the reader, the prevalence of consolidation, crazy-paving pattern, and OP pattern ranged 66–80%, 22–66%, and 50–66%, respectively. Most HRCTs were comparably classified as having an RSNA "typical appearance" by CR (76%), RR1 (86%), and RR2 (68%).

3.2. Agreement between RR1-CR and RR2-CR on disease extent, HRCT signs, and RSNA categorization

Results are summarized in Table 3. The inter-reader agreement we observed in assessing COVID-19 pneumonia disease extent at HRCT was almost perfect both between RR1-CR (k = 0.83, $p < 0.001$) and RR2-CR (k = 0.88, p < 0.001). The agreement on consolidation, crazy paving pattern, OP pattern, and pulmonary artery diameter was substantial, moderate, slight, and good-to-excellent, respectively. The agreement in classifying the HRCTs as per the RSNA categorization was moderate for RR1 versus CR (k = 0.56) and substantial for RR2 versus CR (k = 0.67).

Both RR1 and RR2 showed a higher agreement with CR in assessing disease extent in negative-to-mild cases (almost perfect agreement, with PABAK = 0.90 between RR1-CR and RR2-CR) as compared to moderate-to-severe ones (moderate to substantial agreement, with k = 0.50 between RR1-CR, and k = 0.61 between RR2-CR). The inter-reader agreement ranged fair to almost perfect when assessing different HRCT variables in negative-to-mild cases and none to almost perfect in moderate-to-severe cases (Supplementary material 3), suggesting higher variability and overall lower agreement in the latter category of patients.

Table 1
Patient-related clinical characteristics.

Clinical variable	All patients (n = 150)
Age	Mean ± S.D.
	64.6 ± 14 years
Sex	n (%)
Female	53 (35)
Male	97 (65)
Comorbidities, number	n (%)
0	45 (30)
1	51 (34)
2	40 (27)
3	11 (7)
4	3 (2)
Comorbidities, type	n (%)
Cardiovascular	82 (55)
Respiratory	14 (9)
Chronic renal failure	10 (7)
Tumors	23 (15)
Obesity	17 (11)
Diabetes mellitus	28 (19)
Immunocompromise	2 (1)
Type of hospitalization	n (%)
No hospital admission	43 (29)
Hospitalization without ICU admission	78 (52)
Hospitalization including ICU admission	29 (19)

Table 2
Distribution of disease extent, prevalence of HRCT signs associated with unfavorable prognosis, and RSNA categorization on the test pool (n = 50).

	RR1[a]	RR2	CR[b]
Disease extent			
Continuous variable, *CT involvement score*			
Mean (SD)	8.42 (5.75)	7.36 (4.88)	7.76 (5.02)
Median (range)	7.00 (0–23)	6.00 (0–22)	6.50 (0–23)
Categorical variable, n (%)			
Negative/mild	21 (42)	21 (42)	22 (44)
Moderate	15 (30)	17 (34)	20 (40)
Severe	14 (28)	12 (24)	8 (16)
HRCT[c] signs			
Categorical variable, n (%)			
Consolidation	34 (68)	40 (80)	33 (66)
Crazy-paving pattern	33 (66)	11 (22)	17 (34)
Organizing pneumonia pattern	27 (54)	25 (50)	33 (66)
Continuous variable, median (range)			
Pulmonary artery diameter	27.00 (20–36)	27.00 (21–37)	27.00 (20–37)
RSNA[d] categorization			
Categorical variable, n (%)			
Typical	43 (86)	34 (68)	38 (76)
Atypical	5 (10)	6 (12)	8 (16)
Indeterminate	1 (2)	6 (12)	1 (2)
Negative	1 (2)	4 (8)	3 (6)

Footnotes:
[a] Radiology resident 1.
[b] Chest-devoted radiologist.
[c] High-resolution computed tomography.
[d] Radiological Society of North America.

Two example cases are illustrated in Figs. 3 and 4.

3.3. Prediction of patient's clinical outcome

Based on the clinical utility analysis performed in the whole set of 150 patients' readings by CR, we found that disease extent, consolidation, crazy-paving pattern, and OP pattern were significantly associated with hospitalization and ICU admission, independently of the presence and number of patient comorbidities (Table 4). The highest odds ratio (OR) was observed for severe disease extent at HRCT for both hospitalization (OR = 55) and ICU admission (OR = 49).

PA diameter was an independent predictor of ICU admission but not for hospitalization.

4. Discussion

In this study, we showed that at the end of a dedicated 12 days-long crash course, readers with no experience in thoracic imaging were (i) comparable to an experienced thoracic radiologist in assessing the HRCT extent of COVID-19 (almost perfect agreement) and categorizing pulmonary involvement as typical for COVID-19 according to RSNA categorization (moderate-to-substantial agreement); (ii) able to acceptably report most other HRCT features known to be associated to adverse outcomes.[4–10] Our results suggest that HRCT-naïve readers can be readily made available for focused clinical activity in a non-standard scenario with a high prevalence of disease, i.e., to face the high burden of HRCTs required to support prompt clinical decision-making in a pandemic or epidemic. This goal might be achieved by attending the course when a pandemic/epidemic presents or between different waves.

When building the crash course, we aimed to prioritize the tasks to be learned by participants. We assumed that non-subspecialty radiologists' role would reasonably consist of replying to key clinical questions rather than providing refined judgment of subtler findings. In this light, we aimed primarily to evaluate whether inexperienced readers could reliably assess lung disease extent, as this feature was shown to predict severe clinical outcomes in previous coronavirus-related epidemics,[26,27] current COVID-19 pandemic,[9,10,28] and pulmonary infections in general.[29] When asked to report examinations of the test pool, mimicking clinical activity, RR1 and RR2 showed almost perfect agreement with CR in assessing the CTIS, which in turn was an independent predictor of patient hospitalization and ICU admission. The choice of disease extent as the primary goal can be further supported by previous results suggesting that this feature is associated with substantial-to-excellent inter-reader agreement between readers of different experience,[30,31] although no previous studies involved HRCT-naïve readers, as far as we know. Of note, we involved HRCT-naïve residents with still limited experience in radiology to minimize the confounding effects of previous experience and expertise. The inter-reader agreement in evaluating secondary HRCT features was quite variable in the case of more severe disease. This was an expected result, as interpreting those findings reasonably requires advanced expertise to be consistently reproduced. However, while lower than negative-to-mild disease, the inter-reader agreement in assessing disease extent was moderate to substantial, suggesting that the crash course could reliably prepare readers for the main clinical goal of assessing disease severity in complex HRCT patterns.

In patients with COVID-19 pneumonia, consolidation and crazy-paving pattern were hypothesized to reflect a combination of alveolar edema and damage, interstitial inflammatory changes, and possible bacterial superinfection resulting from immune system activation triggered by the virus itself.[9,32] Proper identification of these signs is of clinical importance since their demonstration early in the course of the disease is associated with severe pneumonia.[7] In addition, consolidation

Table 3
Inter-reader agreement between RR1-CR and RR2-CR on disease extent, HRCT signs and RSNA categorization on the test pool (n = 50).

HRCT[a] variable	RR1[b]-CR[d]			RR2[c]-CR[d]		
	k (95%C.I.)	Agreement level	p-Value	k (95%C.I.)	Agreement level	p-Value
Disease extent	0.83 (0.72–0.94)	Almost perfect	<0.001	0.88 (0.73–0.96)	Almost perfect	<0.001
Consolidation	0.68 (0.44–0.87)	Substantial	<0.001	0.65 (0.41–0.87)	Substantial	<0.001
Crazy-paving pattern	0.42 (0.24–0.63)	Moderate	<0.001	0.51 (0.22–0.75)	Moderate	<0.001
Organizing pneumonia pattern	0.10 (−0.17–0.37)	Slight	0.48	0.20 (−0.06–0.45)	Slight	0.14
Pulmonary artery diameter	0.87 (0.80–0.92)[f]	Good	<0.001	0.91 (0.85–0.94)[f]	Excellent	<0.001
RSNA[e] categorization	0.56 (0.22–0.82)	Moderate	<0.001	0.67 (0.41–0.84)	Substantial	<0.001

Footnotes:
[a] High-resolution computed tomography.
[b] Radiology resident 1.
[c] Radiology resident 2.
[d] Chest-devoted radiologist.
[e] Radiological Society of North America.
[f] Intraclass correlation coefficient, ICC (95%C.I.)

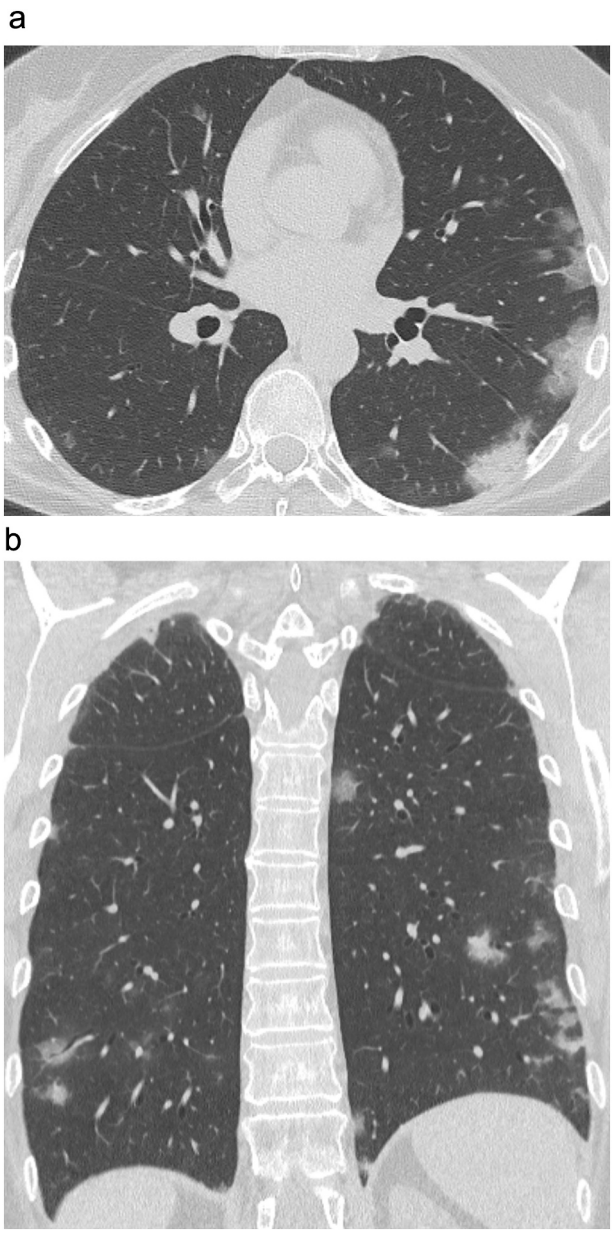


Fig. 3. COVID-19 pneumonia with mild disease extent in a 60-year-old female patient. HRCT images on the axial (a) and coronal (b) planes show peripheral, multifocal, bilateral ground-glass opacities, mainly involving the lower lobes. Computed tomography involvement scores were 5/24 for radiology resident 1, 5/24 for radiology resident 2, and 4/24 for the chest-devoted radiologist. All the readers defined the appearance as "typical" according to the Radiological Society of North America categorization. The patient was discharged from the hospital the same day as the HRCT was made.

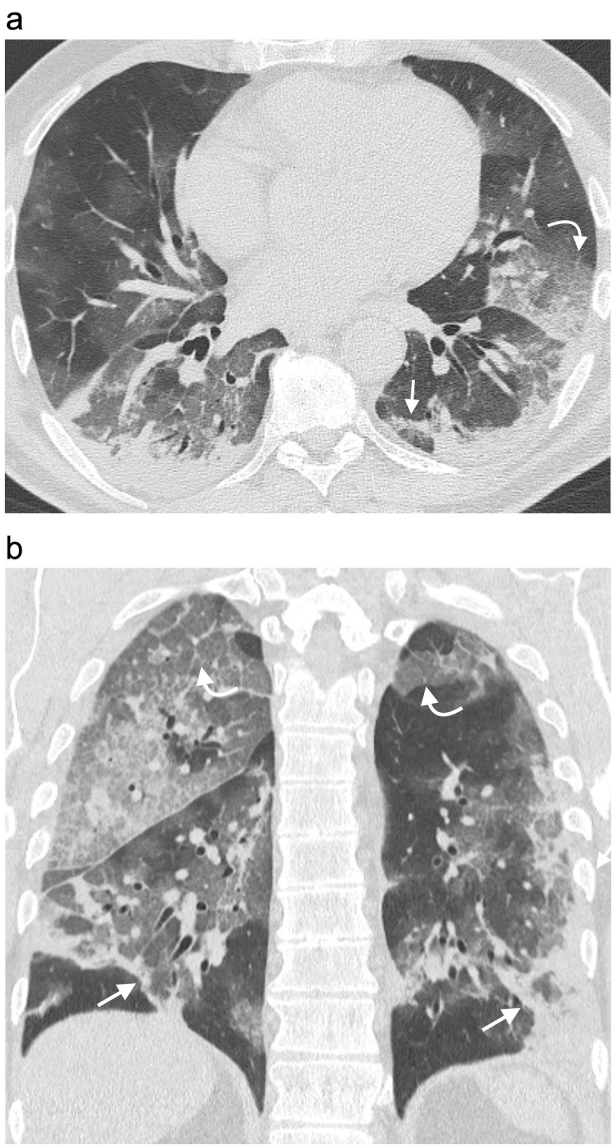


Fig. 4. COVID-19 pneumonia with severe disease extent in a 62-year-old male patient. HRCT images on the axial (a) and coronal (b) planes show bilateral, multifocal ground-glass opacities, along with subpleural consolidations in the lower lobes and areas of crazy-paving pattern in the upper lobes (*curved arrows*). Perilobular consolidations and band-like opacities are also visible, resembling an organizing pneumonia pattern (*straight arrows*). Computed tomography involvement scores were 16/24 for radiology resident 1, 17/24 for radiology resident 2, and 21/24 for the chest-devoted radiologist. All the readers defined the appearance as "typical" according to the Radiological Society of North America categorization. The patient was hospitalized and admitted to the intensive care unit.

was shown to be a common early finding in other coronavirus-related pneumonias,[33] suggesting it could be reasonably frequent even in future viral pandemics/epidemics. We observed that the agreement between RR1-RR2 versus CR was substantial for consolidation and moderate for crazy-paving pattern. Of note, these levels of agreement mirrored what was previously found in the COVID-19 setting[30,34,35] and non-COVID-19 setting[36] when comparing radiologists with different experience. This result suggests that a crash course has the potential to provide a minimum level of inter-reader agreement lying within expected variability for these relevant signs.

Concerning the remaining HRCT signs, we observed good-to-excellent inter-reader agreement for PA diameter. According to a recent consensus proposal,[12] this finding should always be reported, as a larger diameter (e.g., >29 mm) was found to predict a worse prognosis.[4,5] On the other hand, the course appeared inadequate in achieving acceptable agreement in the case of OP pattern, as somewhat expected from the difficulty inexperienced readers can have in meeting its complex definition.[20,37] This is supported by a recent paper showing great inter-reader variability in assessing different HRCT signs underlying OP pattern in COVID-19 pneumonia, e.g., rhomboid shape, band-like pattern, reversed halo sign, with k values ranging from 0.38 to 0.96 depending on the feature.[38] Future iterations of our crash course might include a more specific focus on OP pattern to increase the ability to assess it.

Our study shows some limitations. First, we did not measure the

Table 4
Results from the logistic regression models for hospitalization and intensive care unit admission on the whole set of patients' readings by CR ($n = 150$).

		Univariable analysis		Multivariable analysis			
				Adjusted by comorbidities, number		Adjusted by comorbidities, presence/absence	
		OR[c] (95%C.I.)	p-Value	OR (95%C.I.)	p-Value	OR (95%C.I.)	p-Value
Hospitalization versus no hospital admission							
Disease extent							
Continuous variable, CTIS[b]		1.30 (1.18–1.42)	<0.001	1.31 (1.18–1.44)	<0.001	1.31 (1.19, 1.45)	<0.001
Categorical variable	Moderate vs. negative/mild	5.11 (2.20–11.86)	<0.001	5.2 (2.21–12.19)	<0.001	5.2 (2.22–12.18)	<0.001
	Severe vs. negative/mild	54.55 (6.96–427.59)	<0.001	53.27 (6.77–419.09)	<0.001	56.34 (7.15–444.29)	<0.001
Consolidation (1 vs. 0)		6.74 (1.53–29.74)	0.012	6.56 (1.47–29.18)	0.014	6.8 (1.53–30.23)	0.012
Crazy-paving pattern (1 vs. 0)		6.15 (2.56–14.80)	<0.001	6.32 (2.56–15.58)	<0.001	5.92 (2.45–14.33)	<0.001
Organizing pneumonia pattern (1 vs. 0)		3.22 (1.15–9.00)	0.026	3.65 (1.27–10.48)	0.016	3.91 (1.35–11.30)	0.012
Pulmonary artery diameter (continuous)		1.12 (1.01–1.24)	0.028	1.1 (0.99–1.22)	0.078	1.11 (1.00–1.23)	0.055
ICU[a] admission versus no hospital admission or non-ICU hospitalization							
Disease extent							
Continuous variable, CTIS[b]		1.35 (1.20–1.52)	<0.001	1.37 (1.21–1.55)	<0.001	1.38 (1.22–1.56)	<0.001
Categorical variable	Moderate vs. negative/mild	9.56 (1.17–78.33)	0.035	8.98 (1.09–74.29)	0.042	9.64 (1.17–79.59)	0.035
	Severe vs. negative/mild	48.57 (6.12–385.5)	<0.001	46.8 (5.85–374.64)	<0.001	52.58 (6.51–424.44)	<0.001
Consolidation (1 vs. 0)		7.9 (3.53–17.68)	<0.001	8.22 (3.61–18.73)	<0.001	8.04 (3.56–18.16)	<0.001
Crazy-paving pattern (1 vs. 0)		8.55 (2.86–25.61)	<0.001	8.59 (2.85–25.88)	<0.001	8.32 (2.77–24.96)	<0.001
Organizing pneumonia pattern (1 vs. 0)		3.56 (1.70–7.47)	<0.001	3.97 (1.85–8.55)	<0.001	4.19 (1.92–9.10)	<0.001
Pulmonary artery diameter (continuous)		1.14 (1.02–1.26)	0.019	1.12 (1.01–1.25)	0.035	1.13 (1.01–1.26)	0.028

Footnotes:
[a] Intensive care unit.
[b] Computed tomography involvement score.
[c] Odds ratio.

inter-reader agreement between RR1-RR2 and CR in the training phase, how much it improved over time and after the daily feedback from CR. One might argue that we do not have direct proof that the results on the test phase were determined by what was learned in the training phase. However, it is highly unlikely that inexperienced readers could report test pool HRCTs comparably to CR if not properly trained. On the contrary, any direct measurement during the training phase might have represented an additional source of feedback, thus biasing the agreement towards overestimation. Second, we did not evaluate the post-course performance of RR1 and RR2 in real clinical activity. Indeed, only COVID-19 patients were included in the test set, thus raising the question of whether the agreement we measured was overestimated compared to what is achievable in settings with lower disease prevalence and diagnoses other than COVID-19 pneumonia, or when time occurs between the course and clinical activity. We acknowledge that further studies in a real clinical setting should confirm the efficacy of the crash course and that, if attended in a non-pandemic/epidemic scenario, the course would probably need an on-site or online periodic refresh, as demonstrated in a different clinical setting.[39] Third, the course was not meant to teach readers the main sources of differential diagnosis, such as influenza pneumonia,[40] or identify coexisting diseases. Thus, the crash course cannot prevent readers from missing other relevant findings. However, as discussed above, the objective of a booster approach should be reasonably limited to make it effective in a non-standard scenario rather than a surrogate of the full educational radiology program. Different strategies may overcome this problem, e.g., a delayed quick second opinion by a CR on first-line reports provided by readers who attended a crash course. Fourth, we acknowledge that our sample size might be considered undersized to build a predictive model. However, this was not the study's primary goal, instead focusing on showing the crash course's clinical utility. Our results on outcome prediction match those of previous works,[6–10] suggesting that we achieved our main goal on a reasonably robust and reproducible basis.

In conclusion, we found that at the end of a crash course built on COVID-19 pneumonia cases, HRCT-naïve readers showed almost perfect agreement with an experienced radiologist in assessing disease extent, thus potentially being able to provide relevant clinical information to referring physicians. Inexperienced readers also agreed at a moderate-to-excellent level with CR in providing RSNA categorization and assessing individual HRCT findings associated with prognostic significance (consolidation, crazy-paving pattern, and PA diameter), except for the OP pattern. Our results suggest that, while not representing an instrument for differential diagnosis or subtler evaluation, a crash course can make non-subspeciality radiologists readily available to face a pandemic/epidemic scenario with a high-prevalence of viral pneumonia and HRCT examinations to be reported.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.clinimag.2022.11.010.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. Menni C, Valdes AM, Polidori L, Antonelli M, Penamakuri S, Nogal A, et al. Symptom prevalence, duration, and risk of hospital admission in individuals infected with SARS-CoV-2 during periods of omicron and delta variant dominance: a prospective observational study from the ZOE COVID study. Lancet 2022;399(10335):1618–24. Apr.
2. Kadirvelu B, Burcea G, Quint JK, Costelloe CE, Faisal AA. Variation in global COVID-19 symptoms by geography and by chronic disease: a global survey using the COVID-19 Symptom Mapper. eClinicalMedicine 2022;45:101317. Mar.
3. Kanne JP, Bai H, Bernheim A, Chung M, Haramati LB, Kallmes DF, et al. COVID-19 imaging: what we know now and what remains unknown. Radiology 2021;299(3):E262–79. Jun.
4. Erdoğan M, Öztürk S, Erdöl MA, Kasapkara A, Beşler MS, Kayaaslan B, et al. Prognostic utility of pulmonary artery and ascending aorta diameters derived from computed tomography in COVID-19 patients. Echocardiography 2021;38(9):1543–51. Sep.
5. Zhu QQ, Gong T, Huang GQ, Niu ZF, Yue T, Xu FY, et al. Pulmonary artery trunk enlargement on admission as a predictor of mortality in in-hospital patients with COVID-19. Jpn J Radiol 2021;39(6):589–97. Jun.
6. Li K, Wu J, Wu F, Guo D, Chen L, Fang Z, et al. The clinical and chest CT features associated with severe and critical COVID-19 pneumonia. Invest Radiol 2020;55(6):327–31. Jun.
7. Chon Y, Kim JY, Suh YJ, Lee JY, Park JS, Moon SM, et al. Adverse initial CT findings associated with poor prognosis of coronavirus disease. J Korean Med Sci 2020;35(34):e316.

8. Cereser L, Da Re J, Zuiani C, Girometti R. Chest high-resolution computed tomography is associated to short-time progression to severe disease in patients with COVID-19 pneumonia. Clin Imaging 2021;70:61–6. Feb.
9. Francone M, Iafrate F, Masci GM, Coco S, Cilia F, Manganaro L, et al. Chest CT score in COVID-19 patients: correlation with disease severity and short-term prognosis. Eur Radiol 2020;30(12):6808–17. Dec.
10. Colombi D, Bodini FC, Petrini M, Maffi G, Morelli N, Milanese G, et al. Well-aerated lung on admitting chest CT to predict adverse outcome in COVID-19 pneumonia. Radiology 2020;296(2):E86–96. Aug.
11. Simpson S, Kay FU, Abbara S, Bhalla S, Chung JH, Chung M, et al. Radiological Society of North America expert consensus statement on reporting chest CT findings related to COVID-19. Endorsed by the Society of Thoracic Radiology, the American College of Radiology, and RSNA - secondary publication. J Thorac Imaging 2020;35 (4):219–27. Jul.
12. Neri E, Coppola F, Larici AR, Sverzellati N, Mazzei MA, Sacco P, et al. Structured reporting of chest CT in COVID-19 pneumonia: a consensus proposal. Insights Imaging 2020;11(1):92. Dec.
13. Nordgren Rogberg A, Nyrén S, Westerlund E, Lindholm P. How to train radiology residents to diagnose pulmonary embolism using a dedicated MRI protocol. Acta Radiol Open 2017;6(9):205846011773424. Aug.
14. Wolff L, Su J, Van Loon D, van Es A, van Doormaal PJ, Majoie C, et al. Inter-rater reliability for assessing intracranial collaterals in patients with acute ischemic stroke: comparing 29 raters and an artificial intelligence-based software. Neuroradiology 2022;64(12):2277–84. May 24; Epub ahead of print.
15. Xie X, Zhong Z, Zhao W, Zheng C, Wang F, Liu J. Chest CT for typical coronavirus disease 2019 (COVID-19) pneumonia: relationship to negative RT-PCR testing. Radiology 2020;296(2):E41–5. Aug.
16. Ooi GC, Khong PL, Müller NL, Yiu WC, Zhou LJ, Ho JCM, et al. Severe acute respiratory syndrome: temporal lung changes at thin-section CT in 30 patients. Radiology 2004;230(3):836–44. Mar.
17. Leonard-Lorant I, Severac F, Bilbault P, Muller J, Leyendecker P, Roy C, et al. Normal chest CT in 1091 symptomatic patients with confirmed COVID-19: frequency, characteristics and outcome. Eur Radiol 2021;31(7):5172–7. Jul.
18. Kuzan TY, Murzoğlu Altıntoprak K, Çiftçi HÖ, Ergül U, Ünal Özdemir NB, Bulut M, Yiyit N. A comparison of clinical, laboratory and chest CT findings of laboratory-confirmed and clinically diagnosed COVID-19 patients at first admission. Diagn Interv Radiol 2021 May 7;27(3):336–43.
19. Gotway MB, Reddy GP, Webb WR, Elicker BM, Leung JWT. High-resolution CT of the lung: patterns of disease and differential diagnoses. Radiol Clin North Am 2005; 43(3):513–42. May.
20. Polverosi R, Maffessanti M, Dalpiaz G. Organizing pneumonia: typical and atypical HRCT patterns. Radiol Med 2006;111(2):202–12. Mar.
21. Hansell DM, Bankier AA, MacMahon H, McLoud TC, Müller NL, Remy J. Fleischner society: glossary of terms for thoracic imaging. Radiology 2008;246(3):697–722. Mar.
22. Cicchetti DV, Allison T. A new procedure for assessing reliability of scoring EEG sleep recordings. Am J EEG Technol 1971;11(3):101–10. Sep.
23. Hallgren KA. Computing inter-rater reliability for observational data: an overview and tutorial. Tutor Quant Methods Psychol 2012;8(1):23–34. Feb 1.
24. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33(1):159–74. Mar.
25. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med 2016;15(2):155–63. Jun.
26. Das KM, Lee EY, Enani MA, AlJawder SE, Singh R, Bashir S, et al. CT correlation with outcomes in 15 patients with acute Middle East respiratory syndrome coronavirus. Am J Roentgenol 2015;204(4):736–42. Apr.
27. Ko SF, Lee TY, Huang CC, Cheng YF, Ng SH, Kuo YL, et al. Severe acute respiratory syndrome: prognostic implications of chest radiographic findings in 52 patients. Radiology 2004;233(1):173–81. Oct.
28. Revel MP, Boussouar S, de Margerie-Mellon C, Saab I, Lapotre T, Mompoint D, et al. Study of thoracic CT in COVID-19: the STOIC project. Radiology 2021;301(1): E361–70. Oct.
29. Metlay JP, Waterer GW, Long AC, Anzueto A, Brozek J, Crothers K, et al. Diagnosis and treatment of adults with community-acquired pneumonia. An official clinical practice guideline of the American Thoracic Society and Infectious Diseases Society of America. Am J Respir Crit Care Med 2019;200(7):e45–67. Oct 1.
30. Cereser L, Girometti R, Da Re J, Marchesini F, Como G, Zuiani C. Inter-reader agreement of high-resolution computed tomography findings in patients with COVID-19 pneumonia: a multi-reader study. Radiol Med 2021;126(4):577–84. Apr.
31. Yang R, Li X, Liu H, Zhen Y, Zhang X, Xiong Q, et al. Chest CT severity score: an imaging tool for assessing severe COVID-19. Radiol Cardiothorac Imaging 2020;2 (2):e200047. Apr 1.
32. Tian S, Xiong Y, Liu H, Niu L, Guo J, Liao M, et al. Pathological study of the 2019 novel coronavirus disease (COVID-19) through postmortem core biopsies. Mod Pathol 2020;33(6):1007–14. Jun.
33. Chen X, Zhang G, Hao S, Bai L, Lu J. Similarities and differences of early pulmonary CT features of pneumonia caused by SARS-CoV-2, SARS-CoV and MERS-CoV: comparison based on a systemic review. Chin Med Sci J 2020;35(3):254.
34. Zhang R, Ouyang H, Fu L, Wang S, Han J, Huang K, et al. CT features of SARS-CoV-2 pneumonia according to clinical presentation: a retrospective analysis of 120 consecutive patients from Wuhan city. Eur Radiol 2020;30(8):4417–26. Aug.
35. Sabatino V, Sergio P, Muri M, Zangrandi I, Voltini G, Bosio G, et al. COVID-19: high-resolution computed tomography findings in the first 64 patients admitted to the Hospital of Cremona, the epicentre of the pandemic in Europe. Pol J Radiol 2021;86 (1):172–6.
36. Kumagai S, Arita M, Koyama T, Kumazawa T, Inoue D, Nakagawa A, et al. Prognostic significance of crazy paving ground grass opacities in non-HIV Pneumocystis jirovecii pneumonia: an observational cohort study. BMC Pulm Med 2019;19(1):47. Dec.
37. Baque-Juston M, Pellegrin A, Leroy S, Marquette CH, Padovani B. Organizing pneumonia: what is it? A conceptual approach and pictorial review. Diagn Interv Imaging 2014;95(9):771–7. Sep.
38 Martini K, Blüthgen C, Walter JE, Nguyen-Kim TDL, Thienemann F, Frauenfelder T. Patterns of organizing pneumonia and microinfarcts as surrogate for endothelial disruption and microangiopathic thromboembolic events in patients with coronavirus disease 2019. In: Serra R, editor. PLoS ONE. 15(10); 2020. Oct 5.
39. Hlabangana LT, Andronikou S. Short-term impact of pictorial posters and a crash course on radiographic errors for improving the quality of paediatric chest radiographs in an unsupervised unit — a pilot study for quality-assurance outreach. Pediatr Radiol 2015;45(2):158–65. Feb.
40. Liu M, Zeng W, Wen Y, Zheng Y, Lv F, Xiao K. COVID-19 pneumonia: CT findings of 122 patients and differentiation from influenza pneumonia. Eur Radiol 2020;30(10): 5463–9. Oct.