



Article

# Inter-Observer Agreement between Low-Dose and Standard-Dose CT with Soft and Sharp Convolution Kernels in COVID-19 Pneumonia

Ivan Blokhin <sup>1,\*</sup>, Victor Gombolevskiy <sup>2</sup>, Valeria Chernina <sup>1</sup>, Maxim Gusev <sup>3</sup>, Pavel Gelezhe <sup>1</sup>, Olga Aleshina <sup>4</sup>, Alexander Nikolaev <sup>1</sup>, Nicholas Kulberg <sup>5</sup>, Sergey Morozov <sup>1</sup> and Roman Reshetnikov <sup>1,6</sup>

<sup>1</sup> Research and Practical Clinical Center for Diagnostics and Telemedicine Technologies of the Moscow Health Care Department, 127051 Moscow, Russia; chernina909@gmail.com (V.C.); gelezhe.pavel@gmail.com (P.G.); a.e.nikolaev@yandex.ru (A.N.); morozov@npcmr.ru (S.M.); reshetnikov@fbb.msu.ru (R.R.)

<sup>2</sup> AIRI, 105064 Moscow, Russia; g\_victor@mail.ru

<sup>3</sup> Moscow Polytechnic University, 107023 Moscow, Russia; maxtox3@gmail.com

<sup>4</sup> City Clinical Hospital No. 13 of the Moscow Health Care Department, 115280 Moscow, Russia; olya.aleshina.tula@gmail.com

<sup>5</sup> Federal Research Center “Computer Science and Control” of Russian Academy of Sciences, 119333 Moscow, Russia; kulberg@yandex.ru

<sup>6</sup> Institute of Molecular Medicine, I.M. Sechenov First Moscow State Medical University, 119991 Moscow, Russia

\* Correspondence: i.blokhin@npcmr.ru

**Citation:** Blokhin, I.; Gombolevskiy, V.; Chernina, V.; Gusev, M.; Gelezhe, P.; Aleshina, O.; Nikolaev, A.; Kulberg, N.; Morozov, S.; Reshetnikov, R. Inter-Observer Agreement between Low-Dose and Standard-Dose CT with Soft and Sharp Convolution Kernels in COVID-19 Pneumonia. *J. Clin. Med.* **2022**, *11*, 669. <https://doi.org/10.3390/jcm11030669>

Academic Editors: Francesco Bianco and Nazario Carrabba

Received: 11 December 2021

Accepted: 24 January 2022

Published: 27 January 2022

**Publisher’s Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Abstract:** Computed tomography (CT) has been an essential diagnostic tool during the COVID-19 pandemic. The study aimed to develop an optimal CT protocol in terms of safety and reliability. For this, we assessed the inter-observer agreement between CT and low-dose CT (LDCT) with soft and sharp kernels using a semi-quantitative severity scale in a prospective study (Moscow, Russia). Two consecutive scans with CT and LDCT were performed in a single visit. Reading was performed by ten radiologists with 3–25 years’ experience. The study included 230 patients, and statistical analysis showed LDCT with a sharp kernel as the most reliable protocol (percentage agreement  $74.35 \pm 43.77\%$ ), but its advantage was marginal. There was no significant correlation between radiologists’ experience and average percentage agreement for all four evaluated protocols. Regarding the radiation exposure,  $CTDI_{vol}$  was  $3.6 \pm 0.64$  times lower for LDCT. In conclusion, CT and LDCT with soft and sharp reconstructions are equally reliable for COVID-19 reporting using the “CT 0–4” scale. The LDCT protocol allows for a significant decrease in radiation exposure but may be restricted by body mass index.

**Keywords:** COVID-19; SARS-CoV-2; pneumonia; tomography; X-ray computed

## 1. Introduction

The World Health Organization (WHO) declared a public health emergency on 30 January 2020 due to coronavirus disease 2019 (COVID-19) [1]. While the value of medical imaging in COVID-19 diagnostics remains uncertain [2], computed tomography (CT) is a primary imaging modality used to assess the disease severity in cases of suspected or confirmed COVID-19 pneumonia [3–5].

Studies evaluating CT applicability for COVID-19 indicate the absence of pathognomonic signs but highlight the frequently occurring findings: bilateral peripheral ground-glass opacities and “crazy-paving” pattern with basal predominance [6,7]. The association between the CT findings and COVID-19 depends on the local prevalence of the disease and could be influenced by individual factors such as exposure history [8].

However, CT is widely used for the evaluation of lung parenchyma [9]. Several CT severity score systems for the quantitative assessment of pulmonary lesions in COVID-19 have been described. Yang et al. divided 18 segments of both lungs into 20 regions, subjectively evaluating lung opacities with scores of 0, 1, and 2 for 0%, <50%, or ≥50% involvement of each region, respectively. The final score is the sum of the individual scores in the 20 regions, ranging from 0 to 40 [10]. Salaffi et al. used a system with two scores, “extent” and “nature”, and three evaluation zones for each lung. “Extent” scoring assigns 0 for normal lung, 1 for <25% lung abnormality, 2 for 25–49%, 3 for 50–74% and 4 for ≥75% abnormality. “Nature” scoring reflects the type of pulmonary involvement with 1 for normal lung, 2 for at least 75% ground-glass or crazy paving, 3 for ground-glass, crazy paving and consolidation with less than 75% involvement each, and 4 for at least 75% consolidation; the extent and nature scores are then multiplied by each other and summed at all six levels; the final radiological severity score ranges from 0 to 96 [11]. The latter was initially used by the Russian Federation, but due to its complexity the alternative visual semi-quantitative grading system “CT0-4” was developed and adopted (CT0 category corresponds to 0% lung involvement, CT1: <25%, CT2: 25–50%, CT3: 50–75%, and CT4: >75% involvement) [12–14].

While there is a degree of standardization in assessing COVID-19 pneumonia severity, no normalization is currently applied for the scanning protocol. According to the ALARA principle (As Low As Reasonably Achievable), it is recommended that the lowest possible dose be used, while maintaining the diagnostic quality for the clinical task at hand [15]. WHO issued recommendations on COVID-19 imaging, advocating the usage of a low-dose CT in adults with the disease [16]. Several studies confirm the applicability of low- and ultra-low-dose chest CT protocols for detecting typical COVID-19 pneumonia signs [17,18]. In practice, the effective dose given to COVID-19 patients may vary widely depending on the medical center, and it is generally higher than for patients with other pulmonary infectious diseases [19]. There is also uncertainty about what reconstruction kernel to use for the interpretation of CT images in COVID-19, both for human readers and artificial intelligence. Previous studies demonstrate that kernel selection affects pulmonary nodule measurement in routine [20] and low-dose CT [21], as well as emphysema quantification [22]. A general recommendation for chest CT is to use both soft and sharp kernels [23]. This recommendation also applies to studies in machine learning [24] and radiomics [25]. However, Kwee et al. stated that CT in COVID-19 should be viewed using only a sharp kernel [26].

The study aimed to develop an optimal COVID-19 CT scanning protocol in terms of safety and reliability. For that, we assessed the inter-observer agreement between standard- and low-dose CT with soft and sharp kernels using the CT0-4 severity score.

## 2. Hypotheses

Two hypotheses were put forward to achieve the study objective:

1. The CT0-4 grades assigned by readers for the same patient do not differ between CT and LDCT protocols reconstructed with the sharp or soft kernels;
2. The inter-rater reproducibility of CT0-4 grades assigned by radiologists of different levels of experience does not depend on the kernel and protocol selection.

## 3. Materials and Methods

Ethical approval No. 3/2020 was granted by the regional ethics board, Independent Ethics Committee, Moscow Regional Office of the Russian Society of Radiologists and Radiographers. All participants completed and signed a dedicated consent form before study participation. The study was registered at ClinicalTrials.gov, ID NCT04379531, date of registration 7 May 2020.

### 3.1. Case Selection

This multicenter study was conducted in two outpatient clinics in Moscow, Russia. We included a consecutive sample of patients aged  $\geq 18$  years referred for a chest CT scan by their physician between 6 May 2020, and 22 May 2020, due to suspected viral pneumonia. The clinical symptoms warranting increased suspicion were as follows: increased body temperature ( $>37$  °C); cough without sputum or with a small amount of clear sputum; and increased respiratory rate ( $>22$  breaths per minute). We excluded pregnant and lactating women from the study, as well as patients with implanted foreign objects at the scan level (including cardiac pacemakers and metal constructions); patients with a history of thoracic surgery; cancer patients; and patients refusing to participate in the study.

During a single visit to the CT room, each patient underwent two consecutive scans with standard- and low-dose CT protocols at peak inspiration depth. The scans were acquired consecutively with subjects in the supine position, from the diaphragm to the apex of the lung within a single breath-hold. The patient was not removed from the table between the scans. We did not implement respiratory inductance plethysmography to ensure similar lung volume between the acquisitions due to high patient flow.

### 3.2. CT Acquisition

Chest CT was performed using a 64-detector CT scanner (Aquilion 64, Canon, Japan) without iterative reconstruction algorithms. We used the vendor-provided standard chest CT protocol and had previously developed a low-dose CT protocol for COVID-19 [27].

For both standard- and low-dose CT scans, all data acquisition settings were the same, except for the automatic exposure control (Sure Exposure 3D):

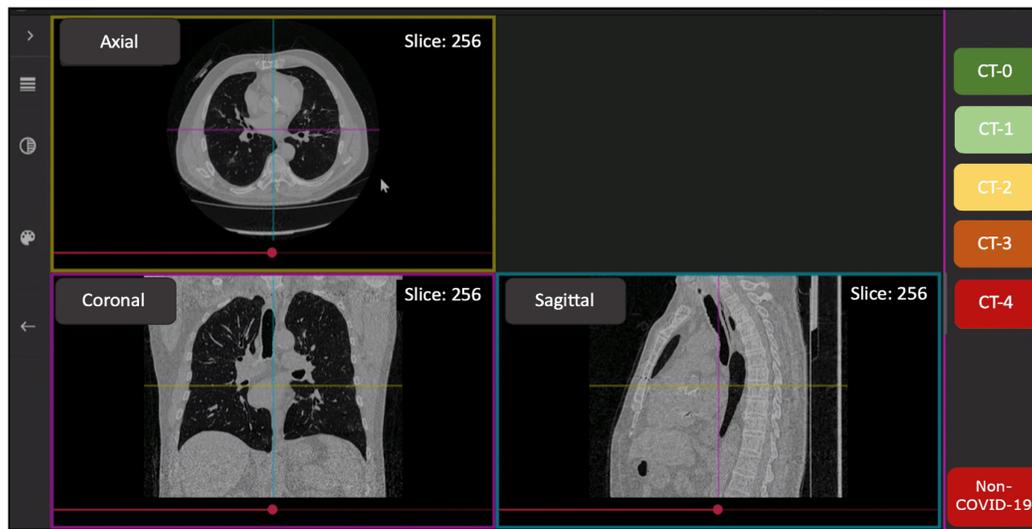
- (1) For a standard CT scan, the tube current was automatically adjusted over the entire scan length from 40 to 500 mA, with a noise level of 10 (standard deviation) for 5.0 mm slices.
- (2) For a low-dose CT scan, the current was automatically adjusted over the entire scan length from 10 to 500 mA, with a noise level of 36 for 5.0 mm slices.

Additional CT scan parameters were the same for standard- and low-dose CT scans: voltage 120 kV; rotation time 0.5 s; direction—out (from legs to head); modulation XY—on; collimation  $64 \times 0.5$  mm; helical pitch 53.0; and scan time—6 s on average. An intravenous contrast agent was not used.

Images were reconstructed as sharp (FC51) and soft (FC07) series with a 1 mm slice thickness for both standard- and low-dose scans.

### 3.3. Image Analysis

Reading of CT scans was performed by ten radiologists with 3–25 years' experience, and trained in COVID-19 pneumonia interpretation. The FAnTom software version 1.2 (Moscow, Russia) [28,29] was modified to assess disease severity according to the CT0-4 scoring scheme (Figure 1); it was also used to facilitate annotation and provide the radiologists with online access to a randomly selected set of anonymized studies.



**Figure 1.** The interface of the FAnTom software tailored for CT0-4 grading.

Each series was independently evaluated by two raters (Reader A and Reader B) blinded to the protocol used (standard- or low-dose CT, sharp or soft kernel) through the FAnTom interface.

### 3.4. Sample Size Consideration

The sample size estimation was made according to Walter et al. [30]. We assumed 0.6 as the minimally acceptable level of correlation between the CT0-4 grades assigned by readers, and 0.7 as a value of expected inter-rater correlation for each of the four protocols. Every case was assessed by two readers; in order to achieve 80% statistical power at a 5% significance level, the study required 205.4 subjects. To account for the multicenter character of the study, we increased this number by 10%, resulting in 230 participants.

### 3.5. Statistical Analysis

For statistical analysis, data from all subjects included in the study (Full Analysis Set, FAS) were used.

The following characteristics were calculated to describe data representing numeric variables (such as age or bodyweight): number of non-missing values; arithmetic mean; and standard deviation.

For qualitative and category variables (sex, CT0-4 grade), the percentage of subjects in each category is provided.

Inter-observer agreement in CT0-4 grades for standard-dose CT with sharp kernel FC51 (CT sharp); standard-dose CT with soft kernel FC07 (CT soft); low-dose CT with sharp kernel FC51 (LDCT sharp); and low-dose CT with soft kernel FC07 (LDCT soft) was analyzed using percent agreement (PA) and Cohen's kappa metrics.

For testing of Hypothesis 1, we compared paired data on CT0-4 scoring for the same patient across the CT sharp, CT soft, LDCT sharp, and LDCT soft protocols. Since for each of the four protocols, there were only two data points per patient (assessments by Reader A and Reader B), the data were unsuitable for Cohen's kappa calculations. For that reason, we tested Hypothesis 1 using only the percentage agreement metric. The data were divided into four matched subsets corresponding to the protocols used, and the PA values between readers A and B were calculated for each case within each subset. These individual agreements were used to calculate the mean PAs for each protocol and compare the protocols using the Wilcoxon signed-rank test [31]. The correlation between the CT0-4 grades assigned by the readers for the same patient was also assessed using Spearman's rank correlation coefficient.

For testing of Hypothesis 2, we collected agreement data for each pair of readers within the four subsets. A pair of readers was included in the analysis if the pair interpreted  $\geq 3$  cases. For every reader, their overall agreement statistic was calculated from pair agreements. These values were then used to calculate mean inter-observer agreement values for each of the four protocols and compare the agreement metrics between the protocols using the Wilcoxon signed-rank test [31].

We also quantified the agreement using the differences between the scores assigned by readers. The difference statistics were compared between the four protocols using the Wilcoxon signed-rank test.

All distributions were checked for normality using a Q–Q plot and the Shapiro–Wilk’s W test [32].

All comparisons were made at a statistical significance level of 0.05. For statistical analyses, the Stata 14 program (StataCorp, College Station, TX, USA) and dplyr [33], ggplot [34], and irr [35] packages for R 3.6.3 (Vienna, Austria) [36] were used.

#### 4. Results

The total sample size was 230 patients (55.6% females), with the following averages: age  $47 \pm 15$  years, weight  $80 \pm 18$  kg, height  $169 \pm 10$  cm, and body mass index  $27.9 \pm 5.6$  kg/m<sup>2</sup>. The distribution of COVID-19 severity within the sample was 45%, 32%, 17%, 5%, and 1% for CT0, CT1, CT2, CT3, and CT4 categories, respectively.

For the standard CT protocol, the average CT dose index (CTDI<sub>vol</sub>) was  $14 \pm 4$  mGy and the average dose-length product (DLP) was  $591 \pm 192$  mGy\*cm, with an average effective dose of  $10 \pm 3$  mSv.

For the low-dose CT protocol, the average CTDI<sub>vol</sub> was  $4 \pm 3$  mGy and the average DLP was  $184 \pm 118$  mGy\*cm, with an average effective dose of  $3 \pm 2$  mSv.

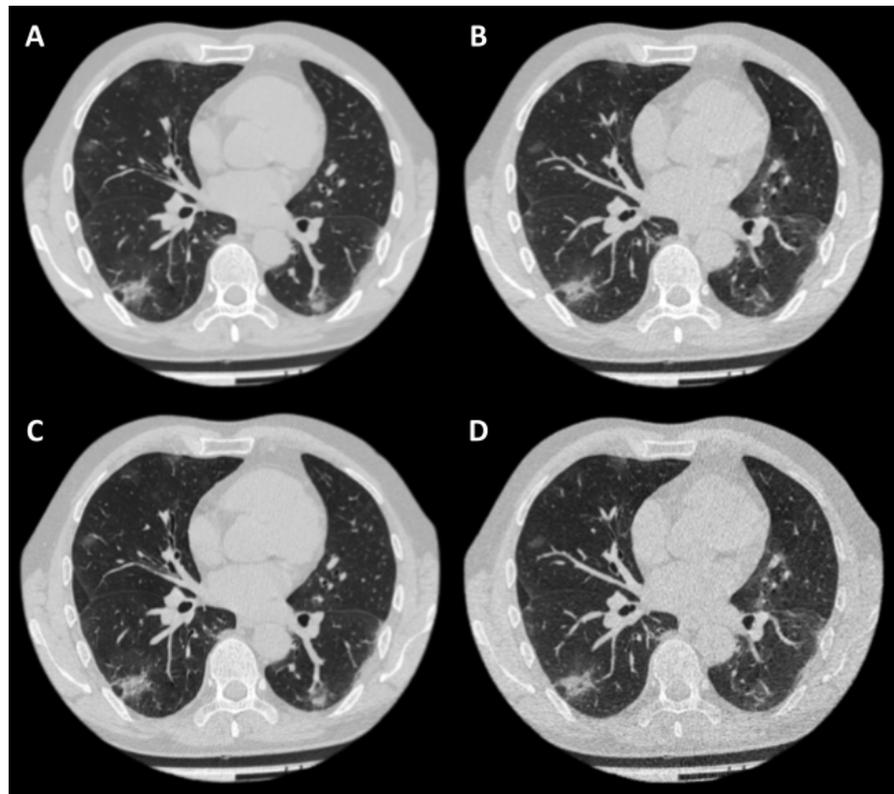
For the low-dose protocol, CTDI<sub>vol</sub> was  $3.6 \pm 0.64$  times lower compared with the standard CT protocol.

Each reader independently interpreted 184 randomly assigned cases. On average, each pair of readers assessed  $7.3 \pm 1.1$  cases of each protocol.

##### 4.1. Hypotheses Findings

**Hypothesis 1 (H1).** The CT0-4 grades assigned by readers to the same patient do not differ between CT and LDCT protocols reconstructed with the sharp or soft kernels (kernel and protocol selection do not affect reporting in COVID-19).

Each case was presented to the readers in one of four reconstructions (Figure 2). The highest inter-observer agreement was observed for the LDCT sharp protocol, while the LDCT soft protocol had the lowest PA value (Table 1). These two protocols were the only pair with a statistically significant difference for which the Wilcoxon rank-sum test rejected Hypothesis 1 (Table 1).



**Figure 2.** Example of a case in four reconstructions. Patient: 57-year-old male, BMI = 24.05 kg/m<sup>2</sup>; (A)—CT soft; (B)—LDCT soft; (C)—CT sharp; (D)—LDCT sharp. CTDI<sub>vol</sub> for CT was 15.5 mGy, for LDCT—3.4 mGy. The effective dose for standard CT was 10.31 mSv, for LDCT—2.46 mSv.

**Table 1.** Inter-observer agreement for matched patient data.

	CT Sharp	CT Soft	LDCT Sharp	LDCT Soft
PA, %	71.74 ± 45.12	66.52 ± 47.29	74.35 ± 43.77	65.65 ± 47.59
Spearman’s rho	0.78, <i>p</i> < 0.001	0.77, <i>p</i> < 0.001	0.82, <i>p</i> < 0.001	0.75, <i>p</i> < 0.001
<i>p</i> -value	0.23	0.07	0.04 *	
		0.53		
			0.84	
			0.16	

\* Statistically significant difference, *p* < 0.05.

**Hypothesis 2 (H2).** The inter-rater reproducibility of CT0-4 grades assigned by radiologists of different levels of experience does not depend on the kernel and protocol selection.

There was no significant correlation between the level of a radiologist’s experience and their average PA value for the CT sharp ( $\rho = -0.01, p = 0.97$ ), CT soft ( $\rho = -0.43, p = 0.22$ ), LDCT sharp ( $\rho = -0.13, p = 0.72$ ), and LDCT soft ( $\rho = -0.24, p = 0.49$ ) protocols. Note that for the soft kernel, the Spearman’s correlation coefficients had lower negative values, suggesting that experienced radiologists were slightly (but not significantly) more likely to disagree with their partner when using this reconstruction method.

Following the results obtained for Hypothesis 1 testing, the protocols using sharp kernels had a higher average inter-rater agreement in comparison with the soft kernel protocols, both in PA and Cohen’s kappa values (Table 2). However, this difference was not statistically significant according to the Wilcoxon test results (Table 2).

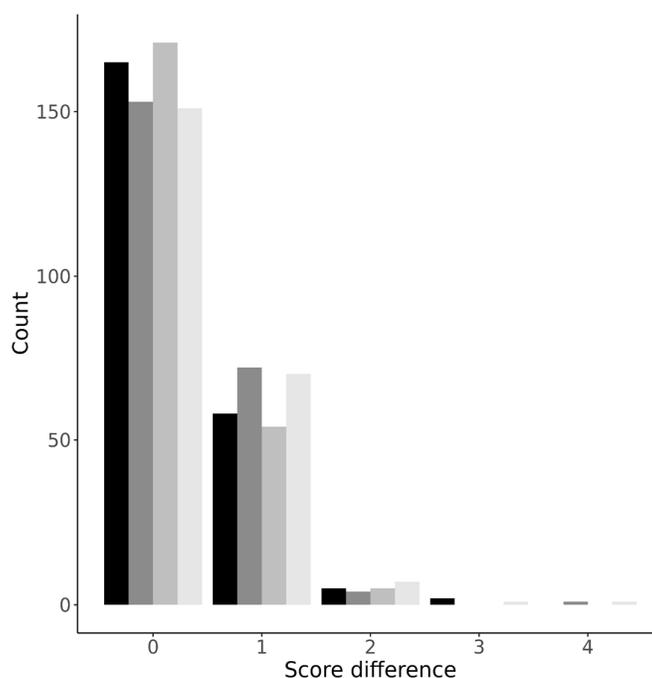
When analyzing the results, we discovered that Cohen’s kappa was not an optimal metric to quantify the inter-rater agreement in CT0-4 grades. For example, there were hard-to-interpret differences between the percent agreement and Cohen’s kappa values for some pairs of readers (Appendix A). The difference is reflected in the *p*-values when comparing the percentage agreements or Cohen’s kappa values across the protocols (Table 2).

**Table 2.** Average inter-observer agreements for the readers in the study.

Radiologists Experience, Years	CT Sharp		CT Soft		LDCT Sharp		LDCT Soft		
	PA, %	Cohen’s Kappa	PA, %	Cohen’s Kappa	PA, %	Cohen’s Kappa	PA, %	Cohen’s Kappa	
0	3	75.1 ± 13.7	0.59 ± 0.23	82.6 ± 10.8	0.73 ± 0.15	89.3 ± 12.4	0.82 ± 0.21	86.6 ± 13.9	0.79 ± 0.21
1	10	54.0 ± 20.0	0.31 ± 0.29	49.5 ± 14.1	0.29 ± 0.15	63.9 ± 13.1	0.49 ± 0.16	43.6 ± 18.4	0.19 ± 0.19
2	25+	87.8 ± 14.0	0.78 ± 0.26	66.8 ± 31.3	0.55 ± 0.40	82.5 ± 11.2	0.68 ± 0.18	68.2 ± 20.6	0.53 ± 0.28
3	6	70.5 ± 11.1	0.39 ± 0.23	55.5 ± 11.1	0.29 ± 0.20	62.5 ± 21.9	0.38 ± 0.32	59.7 ± 7.5	0.38 ± 0.09
4	2	72.0 ± 18.2	0.49 ± 0.38	74.6 ± 18.7	0.62 ± 0.25	73.4 ± 17.8	0.54 ± 0.3	75.8 ± 19.9	0.65 ± 0.27
5	4	65.6 ± 29.5	0.44 ± 0.50	77.7 ± 15.8	0.55 ± 0.35	62.4 ± 16.5	0.44 ± 0.19	77.1 ± 19.4	0.65 ± 0.27
6	1	74.1 ± 17.4	0.51 ± 0.28	72.0 ± 31.3	0.62 ± 0.40	73.1 ± 16.4	0.56 ± 0.31	54.4 ± 29.2	0.36 ± 0.39
7	7	77.5 ± 6.2	0.64 ± 0.10	64.8 ± 16.2	0.44 ± 0.28	75.4 ± 27.7	0.62 ± 0.42	75.0 ± 18.4	0.61 ± 0.26
8	2	78.4 ± 17.8	0.53 ± 0.38	58.9 ± 11.9	0.44 ± 0.28	82.6 ± 14.4	0.7 ± 0.21	68.5 ± 20.0	0.58 ± 0.29
9	6	73.6 ± 21.1	0.57 ± 0.34	60.5 ± 26.9	0.34 ± 0.35	62.5 ± 17.4	0.44 ± 0.16	63.4 ± 19.1	0.46 ± 0.24
Average		72.9 ± 8.8	0.53 ± 0.13	66.3 ± 10.5	0.49 ± 0.15	72.8 ± 9.8	0.57 ± 0.14	67.3 ± 12.5	0.52 ± 0.17
<i>p</i> -value, PA/Cohen’s kappa		0.19/0.58		0.25/0.31		0.48/0.58			
		0.85/0.68				0.79/0.58			
						0.25/0.79			

**4.2. Differences between the “CT0-4” Scores Assigned by Readers**

To further evaluate the inter-rater agreement, we studied the differences between the CT0-4 grades assigned by the readers. For all four protocols, the difference distribution was non-normal with positive kurtosis. Most differences (69.6 ± 4.2%) were zero, while 27.6 ± 3.8%, 2.3 ± 0.5%, 0.4 ± 0.6%, and 0.2 ± 0.3% of assessments differed by one, two, three, and four grades, respectively (Figure 3).



**Figure 3.** Distribution plot of differences in CT0-4 grades across the four protocols. From darker to lighter: LDCT sharp; LDCT soft; CT sharp; CT soft.

In agreement with the results obtained for Hypotheses 1 and 2, the most reliable protocol was LDCT sharp, and the protocol with the most mismatches was LDCT soft (Figure 3). While the difference between the protocols was not statistically significant, we find this observation interesting as it highlights again the marginal advantage of the LDCT sharp protocol over the alternatives.

## 5. Discussion

We assessed the inter-observer agreement between standard- and low-dose CT with soft and sharp kernels using the CT0-4 severity score. The most reliable protocol in terms of percentage agreement was LDCT reconstructed with a sharp kernel, and the protocol with the most mismatches was LDCT reconstructed with a soft kernel. These two protocols were the only pair with statistically significant differences for inter-observer agreement. There was no significant correlation between the level of radiologists' experience and their average percentage agreement value for all four evaluated protocols. Regarding the radiation exposure,  $CTDI_{vol}$  was  $3.6 \pm 0.64$  times lower for the low-dose protocol, compared with the standard CT protocol.

The proposed LDCT protocol for COVID-19 reconstructed with soft and sharp kernels is associated with a significant decrease in the effective dose and is unlikely to negatively impact the accuracy of LDCT reporting in COVID-19. A similar conclusion was reached by Samir et al. in a prospective study comparing ultra-low-dose CT with a 22 mAs tube current and low-dose CT. The former had excellent inter-observer agreement (96–100%) and 90.38–93.84% accuracy in COVID-19 imaging [18]. Tabatabaei et al. obtained similar results on a small sample of 20 patients with an intraclass correlation of 0.98–0.99 for detecting COVID-19 pneumonia using a 30 mAs low-dose chest CT protocol (effective dose of  $1.80 \pm 0.42$  mSv) [17]. Note that both studies used binary classification for detecting signs of pneumonia, but without severity assessment. To our knowledge, no prior studies have compared the standard- and low-dose CT protocols with different convolution kernels using a semi-quantitative scale for lung lesions in suspected COVID-19.

Regarding the convolution kernel choice, our findings agree with the recommendations of Kwee et al., suggesting usage of the sharp kernel for pulmonary tissue evaluation in COVID-19 patients [26]. The marginal advantage of the combination of the low-dose protocol and sharp convolution kernel is also consistent with the Fleischner Society guidelines for pulmonary nodule management [20,37].

Image noise is one of the main factors determining the diagnostic performance of CT, as it has a direct effect on image quality and, therefore, on the radiation exposure level specified for a study. Overweight and obese patients ( $BMI \geq 25$ ) often suffer from reduced CT image quality at a typical effective dose due to higher noise levels. It is particularly pertinent in chest CT, especially when implementing specialized scanning protocols [38,39].

The data on chest LDCT in overweight patients are contradictory. For example, according to Ohana et al., LDCT has limited potential in comparison with standard CT in obese patients ( $BMI > 35$ ) with interstitial pneumonia [40]. These findings restrict the applicability of the *LDCT sharp* protocol to the subjects with a normal BMI.

In addition to gold-standard imaging (CT and LDCT), chest radiography (CXR), lung ultrasound (LUS), and magnetic resonance imaging (MRI) can also be used for COVID-19-associated pneumonia detection [41,42]. LUS has emerged as a radiation-free alternative to CT and LDCT, but its practical applicability is controversial. Some studies show a high level of diagnostic agreement between LUS and CT [43], while others point to its subjectivity and operator dependence [44], restricting LUS usage to critically ill patients, pregnant women, children, and the bedridden aged population [45]. CXR is readily available and operator independent. Ghosh et al. suggest CXR as the initial imaging modality in every suspected case of COVID-19 regardless of laboratory status [42]. Finally, several studies show chest MRI to be promising in COVID-19-associated pneumonia diagnosis and severity assessment [46,47], but its limitations include poor availability, longer examination times, and high susceptibility of the MRI equipment to contamination [48].

Our study has several limitations. First, the study was performed on one CT scanner model, as it is the most common device in the equipment parks at Moscow City Health Department medical organizations. The recommended protocols for other models and manufacturers are likely to differ from that proposed. Second, we did not compare the regional CT0-4 system with other generally accepted semi-quantitative scales [10,11,49]. Third, we did not evaluate the feasibility of noise suppression with iterative reconstructions, neural networks or X-ray beam filtration. The use of a tin filter is not yet widespread but allows a nearly 90% effective dose reduction in chest CT in suspected COVID-19 by filtering the X-ray beam [50]. The increasing prevalence of units with iterative reconstruction or a tin filter may serve as a topic for future research. Fourth, we did not implement respiratory inductance plethysmography to ensure similar lung volume between the acquisitions.

## 6. Conclusions

Standard CT and LDCT protocols with soft and sharp image reconstructions are equally reliable for COVID-19 reporting using the CT0-4 grading system with a marginal advantage of the *LDCT sharp* protocol. The LDCT protocol allows for a significant decrease in radiation exposure, but its applicability may be restricted by the patient's body mass index.

**Author Contributions:** Conceptualization, V.G. and V.C.; data curation, I.B.; formal analysis, R.R.; investigation, I.B., V.G., V.C., P.G., O.A., and A.N.; methodology, V.G.; project administration, V.G.; resources, M.G. and S.M.; software, M.G.; supervision, N.K. and S.M.; validation, I.B.; writing—original draft, I.B., V.C., and P.G.; writing—review and editing, I.B. and R.R. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Institutional Review Board of Moscow Regional Office of the Russian Society of Radiologists and Radiographers (No. 3/2020).

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the study.

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author. The data are not publicly available due to regional legislative policies.

**Acknowledgments:** The authors would like to express their gratitude to Sokolina I.A., Idrisov I.M., Lepikhina D.N., Semenyukov A.A., Trebushenkova I.Y., Nadelaeva I.I., Klyashtorny V.G., Boiko D.A., Laipan A.S., Elizarov A.B., and Gonchar A.P. for their help and input, which are much appreciated.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Appendix A

**Table A1.** Example of differences between percent agreement and Cohen's kappa metrics.

Patient Number	CT0-4 Grade, Reader 0	CT0-4 Grade, Reader 5	Patient Number	CT0-4 Grade, Reader 0	CT0-4 Grade, Reader 6
14	2	2	10	1	2
34	2	2	11	0	0
76	1	1	13	0	0
100	1	2	28	0	0
107	0	0	59	0	0
120	1	0	181	1	2
174	1	0	212	1	0
182	0	0	220	0	0
219	0	0	--	--	--
PA, %	66.7		62.5		
Cohen's kappa	0.518		0.294		

Reader 0 interpreted 9 and 8 cases with readers 5 and 6, respectively. The readers were close in their assessments according to the percent agreement values (Table A1). However, Cohen's kappa values within the pairs differed almost two-fold.

## References

- Lai, C.-C.; Shih, T.-P.; Ko, W.-C.; Tang, H.-J.; Hsueh, P.-R. Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) and Coronavirus Disease-2019 (COVID-19): The Epidemic and the Challenges. *Int. J. Antimicrob. Agents* **2020**, *55*, 105924. <https://doi.org/10.1016/j.ijantimicag.2020.105924>.
- Islam, N.; Ebrahimzadeh, S.; Salameh, J.P.; Kazi, S.; Fabiano, N.; Treanor, L.; Absi, M.; Hallgrimson, Z.; Leeftang, M.M.; Hooft, L.; et al. Thoracic Imaging Tests for the Diagnosis of COVID-19. Edited by Cochrane Infectious Diseases Group. *Cochrane Database Syst. Rev.* **2021**. <https://doi.org/10.1002/14651858.CD013639.pub4>.
- Aljondi, R.; Alghamdi, S. Diagnostic Value of Imaging Modalities for COVID-19: Scoping Review. *J. Med. Internet Res.* **2020**, *22*, e19673. <https://doi.org/10.2196/19673>.
- Harahwa, T.A.; Yau, T.H.; Lim-Cooke, M.S.; Al-Haddi, S.; Zeinah, M.; Harky, A. The Optimal Diagnostic Methods for COVID-19. *Diagnosis* **2020**, *7*, 349–356. <https://doi.org/10.1515/dx-2020-0058>.
- Blažič, I.; Brkljačić, B.; Frija, G. The Use of Imaging in COVID-19—Results of a Global Survey by the International Society of Radiology. *Eur. Radiol.* **2021**, *31*, 1185–1193. <https://doi.org/10.1007/s00330-020-07252-3>.
- Silverstein, W.K.; Stroud, L.; Cleghorn, G.E.; Leis, J.A. First Imported Case of 2019 Novel Coronavirus in Canada, Presenting as Mild Pneumonia. *Lancet* **2020**, *395*, 734. [https://doi.org/10.1016/S0140-6736\(20\)30370-6](https://doi.org/10.1016/S0140-6736(20)30370-6).
- Yoon, S.H.; Lee, K.H.; Kim, J.Y.; Lee, Y.K.; Ko, H.; Kim, K.H.; Park, C.M.; Kim, Y.H. Chest Radiographic and CT Findings of the 2019 Novel Coronavirus Disease (COVID-19): Analysis of Nine Patients Treated in Korea. *Korean J. Radiol.* **2020**, *21*, 494. <https://doi.org/10.3348/kjr.2020.0132>.
- Parekh, M.; Donuru, A.; Balasubramanya, R.; Kapur, S. Review of the Chest CT Differential Diagnosis of Ground-Glass Opacities in the COVID Era. *Radiology* **2020**, *297*, E289–E302. <https://doi.org/10.1148/radiol.2020202504>.

9. Pontone, G.; Scafuri, S.; Mancini, M.E.; Agalbato, C.; Guglielmo, M.; Baggiano, A.; Muscogiuri, G.; Fusini, L.; Andreini, D.; Mushtaq, S.; et al. Role of Computed Tomography in COVID-19. *J. Cardiovasc. Comput. Tomogr.* **2021**, *15*, 27–36. <https://doi.org/10.1016/j.jcct.2020.08.013>.
10. Yang, R.; Li, X.; Liu, H.; Zhen, Y.; Zhang, X.; Xiong, Q.; Luo, Y.; Gao, C.; Zeng, W. Chest CT Severity Score: An Imaging Tool for Assessing Severe COVID-19. *Radiol. Cardiothorac. Imaging* **2020**, *2*, e200047. <https://doi.org/10.1148/ryct.2020200047>.
11. Salaffi, F.; Carotti, M.; Tardella, M.; Borgheresi, A.; Agostini, A.; Minorati, D.; Marotto, D.; Di Carlo, M.; Galli, M.; Giovagnoni, A.; et al. The Role of a Chest Computed Tomography Severity Score in Coronavirus Disease 2019 Pneumonia. *Medicine* **2020**, *99*, e22433. <https://doi.org/10.1097/MD.00000000000022433>.
12. Morozov, S.P.; Kuzmina, E.S.; Ledikhova, N.V.; Vladzomyrskyy, A.V.; Trofimenko, I.A.; Mokienco, O.A.; Panina, E.V.; Andreychenko, A.E.; Omelyanskaya, O.V.; Gomboleviskiy, V.A.; et al. Mobilizing the academic and practical potential of diagnostic radiology during the COVID-19 pandemic in Moscow. *Digit. Diagn.* **2020**, *1*, 5–12. <https://doi.org/10.17816/DD51043>.
13. Morozov, S.P.; Reshetnikov, R.V.; Gomboleviskiy, V.A.; Ledikhova, N.V.; Blokhin, I.A.; Mokienco, O.A. Diagnostic accuracy of computed tomography for identifying hospitalizations for patients with COVID-19. *Digit. Diagn.* **2021**, *2*, 5–16. <https://doi.org/10.17816/DD46818>.
14. Qiblawey, Y.; Tahir, A.; Chowdhury, M.E.; Khandakar, A.; Kiranyaz, S.; Rahman, T.; Ibtihaz, N.; Mahmud, S.; Maadeed, S.A.; Musharavati, F.; et al. Detection and Severity Classification of COVID-19 in CT Images Using Deep Learning. *Diagnostics* **2021**, *11*, 893. <https://doi.org/10.3390/diagnostics11050893>.
15. ICRP Publication 103. The 2007 Recommendations of the International Commission on Radiological Protection. *Ann. ICRP* **2008**, *37*, 2–4. Available online: <https://www.icrp.org/publication.asp?id=ICRP%20Publication%20103> (accessed on 8 December 2021).
16. Use of Chest Imaging in COVID-19. Available online: <https://www.who.int/publications-detail-redirect/use-of-chest-imaging-in-covid-19> (accessed on 9 December 2021).
17. Tabatabaei, S.M.; Talari, H.; Gholamrezanezhad, A.; Farhood, B.; Rahimi, H.; Razzaghi, R.; Mehri, N.; Rajebi, H. A Low-Dose Chest CT Protocol for the Diagnosis of COVID-19 Pneumonia: A Prospective Study. *Emerg. Radiol.* **2020**, *27*, 607–615. <https://doi.org/10.1007/s10140-020-01838-6>.
18. Samir, A.; El-Husseiny, R.M.; Sweed, R.A.; Abd El, N.A.; Masoud, M. Ultra-Low-Dose Chest CT Protocol during the Second Wave of COVID-19 Pandemic: A Double-Observer Prospective Study on 250 Patients to Evaluate Its Detection Accuracy. *Egypt. J. Radiol. Nucl. Med.* **2021**, *52*, 136. <https://doi.org/10.1186/s43055-021-00512-2>.
19. Cristofaro, M.; Fusco, N.; Petrone, A.; Albarello, F.; Di Stefano, F.; Pianura, E.; Schininà, V.; Ianniello, S.; Campioni, P. Increased Radiation Dose Exposure in Thoracic Computed Tomography in Patients with COVID-19. *Radiation* **2021**, *1*, 153–161. <https://doi.org/10.3390/radiation1020014>.
20. Bankier, A.A.; MacMahon, H.; Goo, J.M.; Rubin, G.D.; Schaefer-Prokop, C.M.; Naidich, D.P. Recommendations for Measuring Pulmonary Nodules at CT: A Statement from the Fleischner Society. *Radiology* **2017**, *285*, 584–600. <https://doi.org/10.1148/radiol.2017162894>.
21. Sui, X.; Meinel, F.G.; Song, W.; Xu, X.; Wang, Z.; Wang, Y.; Jin, Z.; Chen, J.; Vliegthart, R.; Schoepf, U.J. Detection and Size Measurements of Pulmonary Nodules in Ultra-Low-Dose CT with Iterative Reconstruction Compared to Low Dose CT. *Eur. J. Radiol.* **2016**, *85*, 564–570. <https://doi.org/10.1016/j.ejrad.2015.12.013>.
22. Gierada, D.S.; Bierhals, A.J.; Choong, C.K.; Bartel, S.T.; Ritter, J.H.; Das, N.A.; Hong, C.; Pilgram, T.K.; Bae, K.T.; Whiting, B.R.; et al. Effects of CT Section Thickness and Reconstruction Kernel on Emphysema Quantification: Relationship to the Magnitude of the CT Emphysema Index. *Acad. Radiol.* **2010**, *17*, 146–156. <https://doi.org/10.1016/j.acra.2009.08.007>.
23. Bhalla, A.S.; Das, A.; Naranje, P.; Irodi, A.; Raj, V.; Goyal, A. Imaging Protocols for CT Chest: A Recommendation. *Indian J. Radiol. Imaging* **2019**, *29*, 236–246. [https://doi.org/10.4103/ijri.IJRI\\_34\\_19](https://doi.org/10.4103/ijri.IJRI_34_19).
24. Hoang-Thi, T.N.; Vakalopoulou, M.; Christodoulidis, S.; Paragios, N.; Revel, M.P.; Chassagnon, G. Deep Learning for Lung Disease Segmentation on CT: Which Reconstruction Kernel Should Be Used? *Diagn. Interv. Imaging* **2021**, *102*, 691–695. <https://doi.org/10.1016/j.diii.2021.10.001>.
25. Denzler, S.; Vuong, D.; Bogowicz, M.; Pavic, M.; Frauenfelder, T.; Thierstein, S.; Eboulet, E.I.; Maurer, B.; Schniering, J.; Gabryś, H.S.; et al. Impact of CT Convolution Kernel on Robustness of Radiomic Features for Different Lung Diseases and Tissue Types. *Br. J. Radiol.* **2021**, *94*, 20200947. <https://doi.org/10.1259/bjr.20200947>.
26. Kwee, T.C.; Kwee, R.M. Chest CT in COVID-19: What the Radiologist Needs to Know. *RadioGraphics* **2020**, *40*, 1848–1865. <https://doi.org/10.1148/rg.2020200159>.
27. Gomboleviskiy, V.; Morozov, S.; Chernina, V.; Blokhin, I.; Vassileva, J. A Phantom Study to Optimise the Automatic Tube Current Modulation for Chest CT in COVID-19. *Eur. Radiol. Exp.* **2021**, *5*, 21. <https://doi.org/10.1186/s41747-021-00218-0>.
28. Morozov, S.P.; Gomboleviskiy, V.A.; Elizarov, A.B.; Gusev, M.A.; Novik, V.P.; Prokudaylo, S.B.; Bardin, A.S.; Popov, E.V.; Ledikhova, N.V.; Chernina, V.Y.; et al. A Simplified Cluster Model and a Tool Adapted for Collaborative Labeling of Lung Cancer CT Scans. *Comput. Methods Programs Biomed.* **2021**, *206*, 106111. <https://doi.org/10.1016/j.cmpb.2021.106111>.
29. Kulberg, N.S.; Gusev, M.A.; Reshetnikov, R.V.; Elizarov, A.B.; Novik, V.P.; Prokudaylo, S.B.; Philippovich, Y.N.; Gomboleviskiy, V.A.; Vladzomyrskyy, A.V.; Kamynina, N.N.; et al. Methodology and Tools for Creating Training Samples for Artificial Intelligence Systems for Recognizing Lung Cancer on CT Images. *Health Care Russ. Fed.* **2020**, *64*, 343–350. <https://doi.org/10.46563/0044-197X-2020-64-6-343-350>.
30. Walter, S.D.; Eliasziw, M.; Donner, A. Sample Size and Optimal Designs for Reliability Studies. *Stat. Med.* **1998**, *17*, 101–110. [https://doi.org/10.1002/\(SICI\)1097-0258\(19980115\)17:1<101::AID-SIM727>3.0.CO;2-E](https://doi.org/10.1002/(SICI)1097-0258(19980115)17:1<101::AID-SIM727>3.0.CO;2-E).

31. Wilcoxon, F. Individual comparisons by ranking methods. *Biom. Bull.* **1945**, *1*, 80–83.
32. Shapiro, S.S.; Wilk, M.B. An analysis of variance test for normality (complete samples). *Biometrika* **1965**, *52*, 591–611.
33. Wickham, H.; Francois, R.; Henry, L.; Müller, K. *dplyr: A Grammar of Data Manipulation*. R Package Version 1.0.7. 2021. Available online: <https://CRAN.R-project.org/package=dplyr> (accessed on 8 December 2021).
34. Wickham, H. Elegant graphics for data analysis. *Media* **2009**, *35*, 10–1007.
35. Gamer, M.; Lemon, J.; Singh, I.F.P. *irr: Various Coefficients of Interrater Reliability and Agreement*. R Package Version 0.84.1. 2019. Available online: <https://CRAN.R-project.org/package=irr> (accessed on 8 December 2021).
36. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2020. Available online: <https://www.R-project.org/> (accessed on 8 December 2021).
37. MacMahon, H.; Naidich, D.P.; Goo, J.M.; Lee, K.S.; Leung, A.N.; Mayo, J.R.; Mehta, A.C.; Ohno, Y.; Powell, C.A.; Prokop, M.; et al. Guidelines for Management of Incidental Pulmonary Nodules Detected on CT Images: From the Fleischner Society 2017. *Radiology* **2017**, *284*, 228–243. <https://doi.org/10.1148/radiol.2017161659>.
38. Azadbakht, J.; Khoramian, D.; Lajevardi, Z.S.; Elikaii, F.; Aflatoonian, A.H.; Farhood, B.; Najafi, M.; Bagheri, H. A Review on Chest CT Scanning Parameters Implemented in COVID-19 Patients: Bringing Low-Dose CT Protocols into Play. *Egypt. J. Radiol. Nucl. Med.* **2021**, *52*, 13. <https://doi.org/10.1186/s43055-020-00400-1>.
39. Paul, N.S.; Kashani, H.; Odedra, D.; Ursani, A.; Ray, C.; Rogalla, P. The Influence of Chest Wall Tissue Composition in Determining Image Noise During Cardiac CT. *Am. J. Roentgenol.* **2011**, *197*, 1328–1334. <https://doi.org/10.2214/AJR.11.6816>.
40. Ohana, M.; Ludes, C.; Schaal, M.; Meyer, E.; Jeung, M.-Y.; Labani, A.; Roy, C. Quel avenir pour la radiographie thoracique face au scanner ultra-low dose? *Rev. Pneumol. Clin.* **2017**, *73*, 3–12. <https://doi.org/10.1016/j.pneumo.2016.09.007>.
41. Finance, J.; Zieleskewicz, L.; Habert, P.; Jacquier, A.; Parola, P.; Boussuges, A.; Bregeon, F.; Eldin, C. Low Dose Chest CT and Lung Ultrasound for the Diagnosis and Management of COVID-19. *J. Clin. Med.* **2021**, *10*, 2196. <https://doi.org/10.3390/jcm10102196>.
42. Ghosh, S.; Deshwal, H.; Saeedan, M.B.; Khanna, V.K.; Raoof, S.; Mehta, A.C. Imaging Algorithm for COVID-19: A Practical Approach. *Clin. Imaging* **2021**, *72*, 22–30. <https://doi.org/10.1016/j.clinimag.2020.11.022>.
43. Wang, M.; Luo, X.; Wang, L.; Estill, J.; Lv, M.; Zhu, Y.; Wang, Q.; Xiao, X.; Song, Y.; Lee, M.S.; et al. A Comparison of Lung Ultrasound and Computed Tomography in the Diagnosis of Patients with COVID-19: A Systematic Review and Meta-Analysis. *Diagnostics* **2021**, *11*, 1351. <https://doi.org/10.3390/diagnostics11081351>.
44. Quarato, C.M.; Mirijello, A.; Maggi, M.M.; Borelli, C.; Russo, R.; Lacedonia, D.; Barbaro, M.P.; Scioscia, G.; Tondo, P.; Rea, G.; et al. Lung Ultrasound in the Diagnosis of COVID-19 Pneumonia: Not Always and Not Only What Is COVID-19 “Glitters”. *Front. Med.* **2021**, *8*, 707602. <https://doi.org/10.3389/fmed.2021.707602>.
45. Vetsheva, N.N.; Reshetnikov, R.V.; Leonov, D.V.; Kulberg, N.S.; Mokienko, O.A. Diagnostic Value of Lung Ultrasound in COVID-19: Systematic Review and Meta-Analysis. *Digit. Diagn.* **2020**, *1*, 13–26. <https://doi.org/10.17816/DD46834>.
46. Vasilev, Y.A.; Sergunova, K.A.; Bazhin, A.V.; Masri, A.G.; Vasileva, Y.N.; Semenov, D.S.; Kudryavtsev, N.D.; Panina, O.Y.; Khoruzhaya, A.N.; Zinchenko, V.V.; et al. Chest MRI of Patients with COVID-19. *Magn. Reson. Imaging* **2021**, *79*, 13–19. <https://doi.org/10.1016/j.mri.2021.03.005>.
47. Yang, S.; Zhang, Y.; Shen, J.; Dai, Y.; Ling, Y.; Lu, H.; Zhang, R.; Ding, X.; Qi, H.; Shi, Y.; et al. Clinical Potential of UTE-MRI for Assessing COVID-19: Patient- and Lesion-Based Comparative Analysis. *J. Magn. Reson. Imaging* **2020**, *52*, 397–406. <https://doi.org/10.1002/jmri.27208>.
48. Collins, J.D.; Rowley, H.; Leiner, T.; Reeder, S.; Hood, M.; Dekkers, I.; Tha, K.; Gulani, V.; Kopanoglu, E.; for the ISMRM MR Safety Committee. Magnetic Resonance Imaging during a Pandemic: Recommendations by the ISMRM Safety Committee. *J. Magn. Reson. Imaging* **2021**, *jmri.28006*. <https://doi.org/10.1002/jmri.28006>.
49. Prokop, M.; Van Everdingen, W.; van Rees Vellinga, T.; Quarles van Ufford, H.; Stöger, L.; Beenen, L.; Geurts, B.; Gietema, H.; Krdzalic, J.; Schaefer-Prokop, C.; et al. CO-RADS: A Categorical CT Assessment Scheme for Patients Suspected of Having COVID-19—Definition and Evaluation. *Radiology* **2020**, *296*, E97–E104. <https://doi.org/10.1148/radiol.2020201473>.
50. Kang, Z.; Li, X.; Zhou, S. Recommendation of Low-Dose CT in the Detection and Management of COVID-2019. *Eur. Radiol.* **2020**, *30*, 4356–4357. <https://doi.org/10.1007/s00330-020-06809-6>.