

Article

Comparison of Computed Tomography Scoring Systems in Patients with COVID-19 and Hematological Malignancies

Marta Hałaburda-Rola ^{1,*}, Joanna Drozd-Sokołowska ² , Magdalena Januszewicz ¹ and Laretta Grabowska-Derlatka ¹

¹ IInd Department of Clinical Radiology, Medical University of Warsaw, 01-445 Warsaw, Poland

² Department of Hematology, Transplantation and Internal Diseases, Medical University of Warsaw, 01-445 Warsaw, Poland

* Correspondence: mrola@wum.edu.pl

Simple Summary: COVID-19 pneumonia poses a serious threat in hematologic patients. Computed tomography is an indispensable tool supporting diagnosis. A better and more objective analysis of the extent of pneumonia enables assessment of the extent of the disease as well as the selection of prognostic factors of death. The aim of this study is to compare four different computed tomography scoring systems (three semiquantitative and one qualitative) in hematology patients to better select patients at risk of death and choose the scoring system that is the most feasible for this group of patients.

Abstract: Background: Numerous computed tomography (CT) scales have been proposed to assess lung involvement in COVID-19 pneumonia as well as correlate radiological findings with patient outcomes. Objective: Comparison of different CT scoring systems in terms of time consumption and diagnostic performance in patients with hematological malignancies and COVID-19 infection. Materials and methods: Retrospective analysis included hematological patients with COVID-19 and CT performed within 10 days of diagnosis of infection. CT scans were analyzed in three different semi-quantitative scoring systems, Chest CT Severity Score (CT-SS), Chest CT Score (CT-S), and Total Severity Score (TSS), as well as qualitative modified Total Severity Score (m-TSS). Time consumption and diagnostic performance were analyzed. Results: Fifty hematological patients were included. Based on the ICC values, excellent inter-observer reliability was found among the three semi-quantitative methods with ICC > 0.9 ($p < 0.001$). The inter-observer concordance was at the level of perfect agreement (kappa value = 1) for the mTSS method ($p < 0.001$). The three-receiver operating characteristic (ROC) curves revealed excellent and very good diagnostic accuracy for the three quantitative scoring systems. The AUC values were excellent (0.902), very good (0.899), and very good (0.881) in the CT-SS, CT-S and TSS scoring systems, respectively. Sensitivity showed high levels at 72.7%, 75%, and 65.9%, respectively, and specificity was recorded at 98.2%, 100%, 94.6% for the CT-SS, CT-S, and TSS scoring systems, respectively. Time consumption was the same for Chest CT Severity Score and TSS and was longer for Chest CT Score ($p < 0.001$). Conclusions: Chest CT score and chest CT severity score have very high sensitivity and specificity in terms of diagnostic accuracy. The highest AUC values and the shortest median time of analysis in chest CT severity score indicate this method as preferred for semi-quantitative assessment of chest CT in hematological patients with COVID-19.

Keywords: COVID-19; computed tomography; hematological malignancies; pneumonia; imaging



Citation: Hałaburda-Rola, M.; Drozd-Sokołowska, J.; Januszewicz, M.; Grabowska-Derlatka, L. Comparison of Computed Tomography Scoring Systems in Patients with COVID-19 and Hematological Malignancies. *Cancers* **2023**, *15*, 2417. <https://doi.org/10.3390/cancers15092417>

Academic Editor: Mingyi Chen

Received: 26 February 2023

Revised: 16 April 2023

Accepted: 21 April 2023

Published: 22 April 2023



Copyright: © 2023 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/>).

1. Introduction

In December 2019, the COVID-19 pandemic, caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), began in Wuhan, China, and spread worldwide,

especially affecting the elderly and those with comorbidities [1]. In patients with hematologic malignancies, COVID-19 caused a high mortality rate and often required withdrawal of anticancer treatment [2]. The pandemic has forced healthcare providers to develop dedicated diagnostic procedures. The primary method of diagnosing SARS-CoV-2 virus infection is real-time reverse transcription polymerase chain reaction (RT-PCR) or next-generation sequencing from nasopharyngeal swabs. Imaging techniques are considered additional diagnostic methods performed depending on the severity of clinical symptoms. Chest X-ray has been proven to have no predictive value in patients with COVID-19, as it can give false negatives in patients with mild symptoms [3]. However, in patients admitted to an intensive care unit (ICU), who require oxygen support, mechanical ventilation, sedation, and are in severe clinical conditions, a portable chest X-ray remains an indispensable basic imaging tool [4]. Computed tomography (CT) shows high sensitivity in the diagnosis of suspected COVID-19 and is used to guide patient management [5]. According to the Polish Medical Society of Radiology, the main indications for chest CT are a high risk of progression of COVID-19 or development of its complications. As long as the pulmonary embolism is not suspected, contrast administration is not routinely recommended [4].

Chest CT findings in COVID-19 pneumonia are nonspecific and resemble those in other viral pneumonias, usually presenting as bilateral areas of ground glass opacities (GGO) and consolidations located predominantly peripherally, bilaterally in the lower lobes. Superimposed thickening of the interlobular septum results in a “crazy-paving” appearance. In severe cases, patients present radiological features of acute respiratory distress syndrome (ARDS) [6].

Patients with hematologic diseases usually present with immune deficiencies, including neutropenia either due to the hematological condition itself or anticancer therapy.

Thus, there are often infectious complications and superinfections of the diseases themselves and treatment. Chest CT is an indispensable tool in these cases to assess lung involvement in COVID-19 pneumonia and to evaluate any possible concomitant abnormalities such as nodules with halos or other features suggestive of invasive fungal infections or other pathologies. Chest CT provides an opportunity to objectively assess the degree of lung involvement in patients with COVID-19. Several scales have been developed to evaluate the lesions and enable more objective semi-quantitative and qualitative analysis.

There are a few papers describing the challenges and pitfalls of chest CT imaging in hematologic patients, but there is a lack of information on quantitative CT analysis in this particular group of patients [7]. The aim of this study was to compare the sensitivity and specificity of three semiquantitative CT scoring systems in terms of diagnostic accuracy and time-consumption in patients with hematologic diseases and COVID-19 and interobserver agreement in the qualitative analysis of lung parenchyma.

2. Materials and Methods

2.1. Study Design and Patient Population

This single-center, retrospective study includes 50 consecutive patients with hematological diseases and positive RT-PCR test results for COVID-19, diagnosed between March 2020 and October 2021.

The severity of COVID-19 was assessed according to the recommendations of the Polish Association of Epidemiologists and Infectiologists [8]. According to this classification, stage 1 is asymptomatic or mildly symptomatic, with oxygen saturation as measured by pulse oximetry (SpO₂) of $\geq 94\%$. In stage 1, patients do not require hospitalization because of COVID-19. Stage 2 is fully symptomatic, with SpO₂ $< 94\%$ and the requirement for hospitalization because of the need for oxygen therapy. Stage 3 is defined as respiratory failure with SpO₂ $< 90\%$ and requirement for high-flow oxygen therapy. Stage 4 is defined as ARDS with the need for invasive mechanical ventilation (IMV) and ICU treatment. Stages 1 and 2 were considered as the non-severe COVID-19 group, while stage 3 and 4 were considered as the severe COVID-19 group. The indications for CT were mild, non-specific symptoms of respiratory tract infection and clinical risk factors for developing

complications. CT was performed immediately after a positive SARS-CoV-2 RT-PCR test. Comorbidities, complete blood count parameters, laboratory parameters such as D-dimer, and C-reactive protein (CRP) at the time of admission diagnosis were recorded.

2.2. Chest Computed Tomography

The CT scans were obtained on supine position at the end-inspiratory effort using a GE 64-row scanner. The thoracic scanning parameters were as follows: tube voltage of 120 kV, tube current of 240–460 mA, 1.25 mm collimation, a pitch of 0.984:1, and applying a large field of view and a rotation time of 0.7 s. From this dataset, axial slices of 1 mm were obtained. The coronal and sagittal slices were obtained as high resolution multiplanar reformation images. All CT exams were obtained without intravenous contrast administration. Images were analyzed using picture archiving and communications systems (PACS).

The CT evaluation was performed by two independent, board-certified radiologists. Each observer analyzed the examination in four scales. Time consumption on reporting of each patient scoring system was recorded and calculated in the same reading environment using the same diagnostic monitors.

The software used for the image analysis was GE AW Server 3.2 ext. 4.0.

According to Fleischner Society, a standard glossary of terms for computed tomography was applied to describe the lesions [9]:

- GGO was defined as hazy increased attenuation of lung with preserved bronchial and vascular margins;
- Consolidation was considered an increase in pulmonary parenchymal attenuation that obscures the airways and vessels;
- Crazy paving was the area of GGO with coexisting thickening of interlobular septae;
- Pleural effusion was described as a free fluid in the pleural cavity;
- Septal thickening comprised abnormal widening of an interlobular septum or septae;
- Subpleural lines comprised a thin curvilinear opacity of a few millimeters or less thickness usually less than 1 cm from pleural surface and paralleling the pleura.

The chest CT exams were evaluated according to four scoring systems described below:

- Chest CT Severity Score (CT-SS): According to the anatomic structure, the 18 segments of both lungs were divided into 20 regions, in which the posterior apical segment of the left upper lobe was subdivided into apical and posterior segmental regions, whereas the anteromedial basal segment of the left lower lobe was subdivided into anterior and basal segmental regions. The lung opacities in all of the 20 lung regions were subjectively evaluated on chest CT images using a system attributing score of 0, 1, and 2 if parenchymal opacification involved 0%, less than 50%, or equal to or more than 50% of each region, respectively. The CT-SS was defined as the sum of the individual scores in the 20 lung segment regions, which may range from 0 to 40 points [6];
- Chest CT Score (CT-S): Each of the five lung lobes was visually scored on a scale of 0 to 5, with 0 indicating no involvement; 1, less than 5% involvement; 2, 5–25% involvement; 3, 26–49% involvement; 4, 50–75% involvement; and 5, more than 75% involvement. The total CT score was the sum of the individual lobar scores and ranged from 0 (no involvement) to 25 (maximum involvement) [10,11];
- Total severity score (TSS): The TSS was calculated for each of the 5 lobes in all patients. According to the extent of pulmonary involvement, each lobe could be scored from 0 to 4 points as the following: 0, no involvement; 1, from 1 to 25% involvement; 2, from 26 to 50% involvement; 3, from 51 to 75% involvement; and 4, more than 75% involvement. The sum of each individual lobar score resulted in the TSS, which ranged from 0 to 20 [12];
- Modified Total Severity Score (m-TSS): m-TSS scale includes additional qualitative features of lung involvement: A—ground glass opacity, B—crazy-paving pattern, C—consolidations, and X—character other than enlisted [13].

2.3. Viral Analysis

Nasopharyngeal swab analysis was performed in each patient at admission to the hospital and afterward in patients with clinical symptoms of viral infection or in patients with known contact with SARS-CoV2-infected patients.

The diagnosis of SARS-CoV-2 infection was based on a nucleic acid amplification test (Seegene STARlet[®], Bio-Rad CFX[®], Hologic Panther[®], Warsaw, Poland) using nasopharyngeal swabs.

2.4. Statistical Analysis

The MS Power BI Desktop was applied for finding the outliers via scatter chart in the first phase of the analysis and preliminary statistical analysis. IBM SPSS Statistics (version: 28.0.1.0(142)) was used to analyze the distribution of variables and perform some statistical tests and calculate some statistics. PQStat (version: 1.8.4.) was used for the same purpose to analyze the data and to prepare all the visualizations seen in this article. To determine the inter-observer agreement between the two reviewers, two tests were used. The Kaplan–Meier curve was used to calculate the median overall survival time for ICU cases and PCR test cases. To check the significance of the nominal variables, the test of kappa coefficient was applied. Whereas for the interval variables (after checking and confirming the normal distribution with an appropriate statistical test) a test to check the significance of the ICC coefficient (intraclass correlation coefficient) was used. Intraclass correlation (ICC) was used to examine inter-observer agreement for the three quantitative methods (CT-SS, CT-S, and TSS). The determination of the ROC curve and the calculation of the area under this curve (AUC), as well as the calculation of sensitivity and specificity, were used to compare which of the 3 quantitative methods could generate the best results in assigning patients to the severe and non-severe COVID-19 groups. After comparing the three quantitative methods, i.e., CT-SS, CT-S, and TSS, in terms of inter-observer agreement and in terms of ROC curve, AUC, and sensitivity and specificity values, the time consumptions of the three test methods were compared. For this purpose, the normal distribution of the three variables was checked. When the test showed that one of the variables did not meet the condition of normal distribution (p -value > 0.05) it was decided to perform Friedman’s ANOVA test, which was dedicated to the independent interval variable that did not meet the condition of normal distribution. Patients were divided into two groups: the severe COVID-19 group and non-severe COVID-19 group. To determine the p -value, three different tests were used. The main aim was to define whether the p -value was less than 0.05, indicating a statistically significant difference. For interval scale variables that did not meet the condition of normal distribution and for those that were independent, the Mann–Whitney U test was used, which is specifically designed for such conditions. For quantitative variables that met the condition of normal distribution, the Student’s t -test for independent groups was used. For nominal and independent variables, the chi-square/Fisher’s exact test was used. In order to compare the time consumption of the three semi-quantitative methods of assessing patients’ lungs (Chest CT severity score_time, Chest CT score_time, and Total severity score_time), the normal distribution of these variables was checked at first. Due to the small sample size, the Shapiro–Wilk test was used to test the normal distribution condition of the variables describing the examination/interpretation time in seconds. According to the null hypothesis of the Shapiro–Wilk test, the distribution of the trait under study is a normal distribution. Thus, if the p -value was less than the chosen α level ($\alpha = 0.05$ in this study), the null hypothesis was rejected, and there was evidence that the values did not have a normal distribution. The p -value for this test for the Chest CT severity score time variable was <0.005, i.e., the Chest CT severity score_time variable did not meet the normal distribution condition, whereas the other two variables met the normal distribution condition. For this reason, it was not possible to carry out the test that was previously considered, that is, the single-factor repeated-measures ANOVA test. Instead of the previously considered test, the Friedman ANOVA test, which is dedicated to dependent interval variables meeting the condition of normal distribution, was conducted. Friedman’s

ANOVA test for p -value < 0.005 indicates that not all test times are statistically significantly different from each other.

3. Results

50 patients with hematologic malignancies and COVID-19 diagnosed at the Medical University of Warsaw were included in this study. The clinical data of the patients are presented in Table 1.

Table 1. Clinical and radiological characteristics of the studied patients according to COVID-19 disease progression.

Characteristic	Total	Non-Severe COVID-19	Severe COVID-19	p Value
N	50	28	22	
Age [years]	64.94 ± 17.45	61.32 ± 18.53	69.55 ± 15.16	0.096 ^U
Sex				0.136 ^F
Male	33 (66%)	16 (57.14%)	17 (77.27%)	
Female	17 (34%)	12 (42.86%)	5 (22.73%)	
Mortality rate	20 (40%)	1 (3.57%)	19 (86.36%)	<0.05 ^F
Associated comorbidities	1.92 ± 1.44	1.61 ± 1.2	2.32 ± 1.64	0.1097 ^U
Arterial hypertension	26 (52%)	12 (42.86%)	14 (63.64%)	0.144 ^F
Diabetes	9 (18%)	2 (7.14%)	7 (31.82%)	0.024 ^F
Liver disease	5 (10%)	2 (7.14%)	3 (13.64%)	0.447 ^F
Hyperlipidemia	18 (36%)	9 (32.14%)	9 (40.91%)	0.522 ^F
Heart disease	23 (46%)	14 (50%)	9 (40.91%)	0.522 ^F
Kidney disease	15 (30%)	6 (21.43%)	9 (40.91%)	0.136 ^F
Blood pressure				
Systolic pressure [mmHg]	120.22 ± 18.64	118.86 ± 16.19	121.95 ± 21.64	0.565 ^t
Systolic blood pressure out of the norm	4 (8%)	2 (7.14%)	2 (9.09%)	0.801 ^F
Diastolic pressure [mmHg]	71.16 ± 11.14	72.25 ± 10.51	69.77 ± 12	0.441 ^t
Diastolic blood pressure out of the norm	4 (8%)	2 (7.14%)	2 (9.09%)	0.801 ^F
Admitted to ICU	19 (38%)	3 (10.71%)	16 (72.73%)	<0.05 ^F
Ventilated by a mask with high oxygen volumes	21 (42%)	2 (7.14%)	19 (86.36%)	<0.05 ^F
Intubated	13 (26%)	1 (3.57%)	12 (54.55%)	<0.05 ^F
Saturation [%]	89.04 ± 13.72	94.11 ± 7.96	82.59 ± 16.72	0.001 ^U
D-dimer [μg/L]	2882.42 ± 2771.56	2348.39 ± 2836.64	3562.09 ± 2591.12	0.005 ^U
CRP, C Reactive Protein [mg/L]	141.58 ± 86.98	111.04 ± 79.96	180.45 ± 81.27	0.004 ^t
WBC, White blood cell count [tys/mm ³]	8.69 ± 11.34	9 ± 11.8	8.3 ± 10.99	0.71032 ^U
Neutrophil [tys/μL]	3.9516 ± 5.79	5.17 ± 7.21	2.4 ± 2.61	0.186734 ^U

Test of significance: ^U the Mann–Whitney U test, ^F chi-square or Fisher’s exact test, ^t independent samples t test.

3.1. Clinical and Radiological Characteristics of the Examined Patients According to COVID-19 Severity

In the study sample of patients, 22 (56%) were severe cases, and 28 (44%) belonged to the non-severe COVID-19 group. During the analysis, a statistically significant difference

was noted between the mortality rates of severe and non-severe COVID-19 patients. The Severe COVID-19 group had a mortality rate of 86.36%, while there was only one case of death in the non-severe group, resulting in a mortality rate of 3.57% (p -value < 0.05).

Compared to the non-severe COVID-19 group, patients in the severe COVID-19 group were more likely to demonstrate the presence of a comorbid condition, such as diabetes. Among severely ill COVID-19 patients, 16 (72.73%) were admitted to the ICU, while among non-severe patients, only 3 (10.71%) were admitted to the ICU (p -value < 0.05).

Saturation levels were statistically different between the two groups of patients (Figure 1). The mean saturation level in non-severe patients oscillated at 94.11% (SD 7.96), whereas for severe patients, this value was 82.59% (SD 16.72) (p -value = 0.001). The D-dimer value differed significantly in the groups designated for analysis (Figure 2) between severely ill patients (mean D-dimer 3562.09 $\mu\text{g/L}$; SD 2591.12) and non-severe patients (mean D-dimer 2348.39 $\mu\text{g/L}$; SD 2836.64, p -value = 0.005). The blood-based CRP value (Figure 3) was also higher in severely ill COVID-19 patients (mean CRP 180.45 mg/L; SD 81.27) and non-severe patients (mean CRP 111.04 mg/L; SD 79.96, p -value = 0.004).

The Kaplan–Meier curve was used to determine a curve showing the probability of survival time for the patients from the time of admission to the ICU ($N = 19$), assuming a maximum time in hospital of 15 days. The premise of this analysis was to determine the median time to death after admission to the ICU and to show by means of a curve how the probability of survival time was distributed among the patients studied. While analyzing the Kaplan–Meier curve, it was observed that the median survival from ICU admission was 11 days, with a mean value of slightly over 9 days (Figure 4).

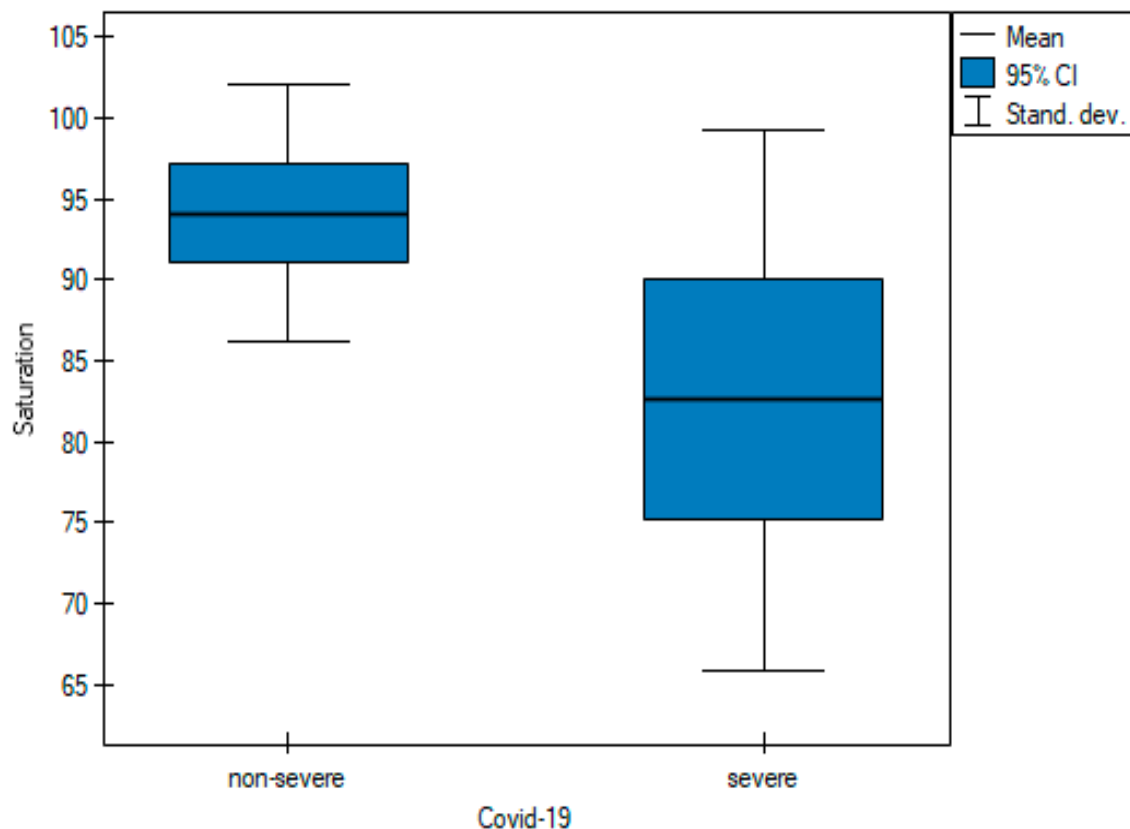


Figure 1. Pairwise comparisons of saturation values of non-severe and severe cases at diagnosis.

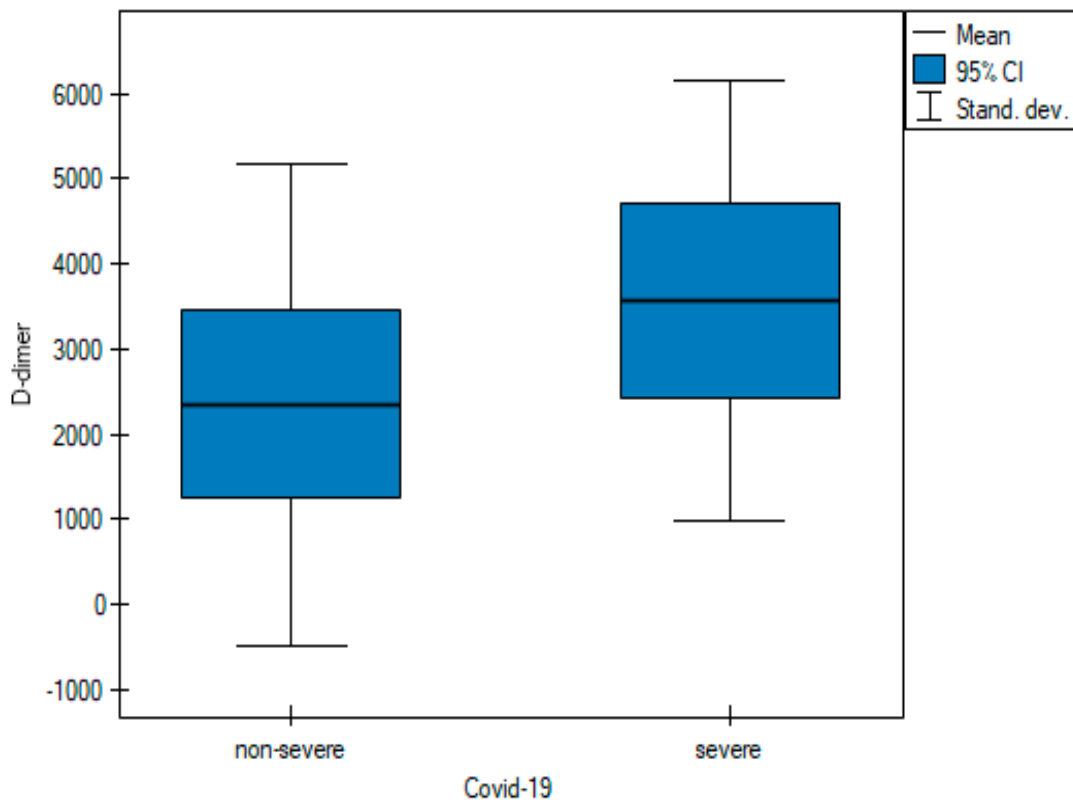


Figure 2. Pairwise comparisons of D-dimer values of non-severe and severe cases at diagnosis.

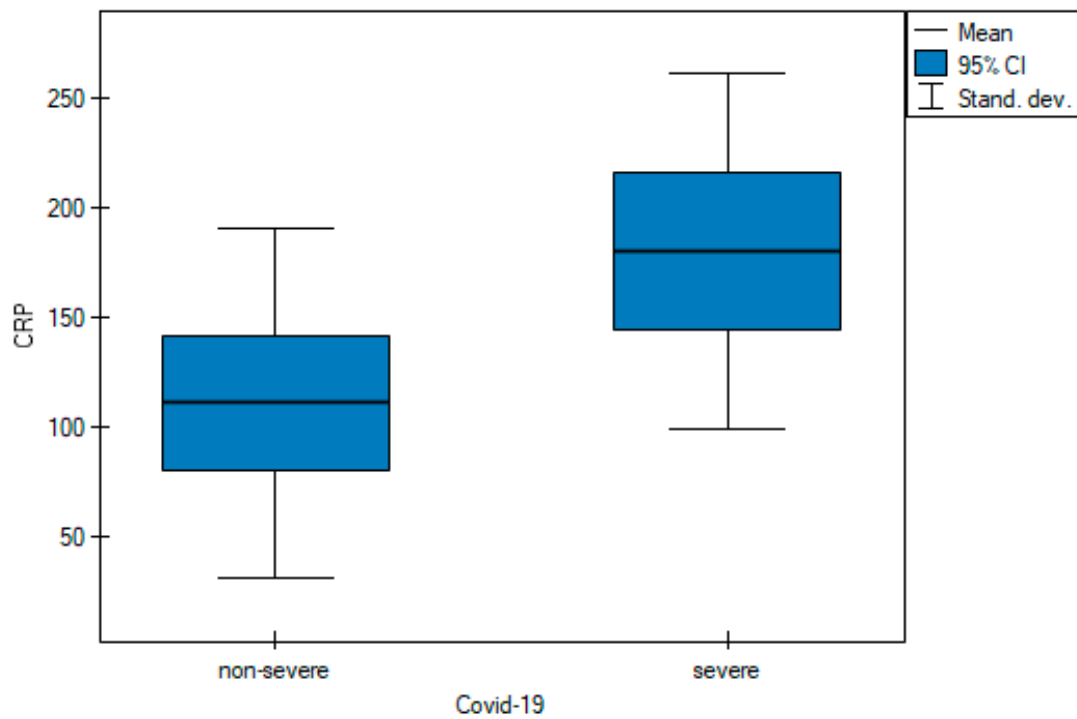


Figure 3. Pairwise comparisons of CRP values of non-severe and severe cases at diagnosis.

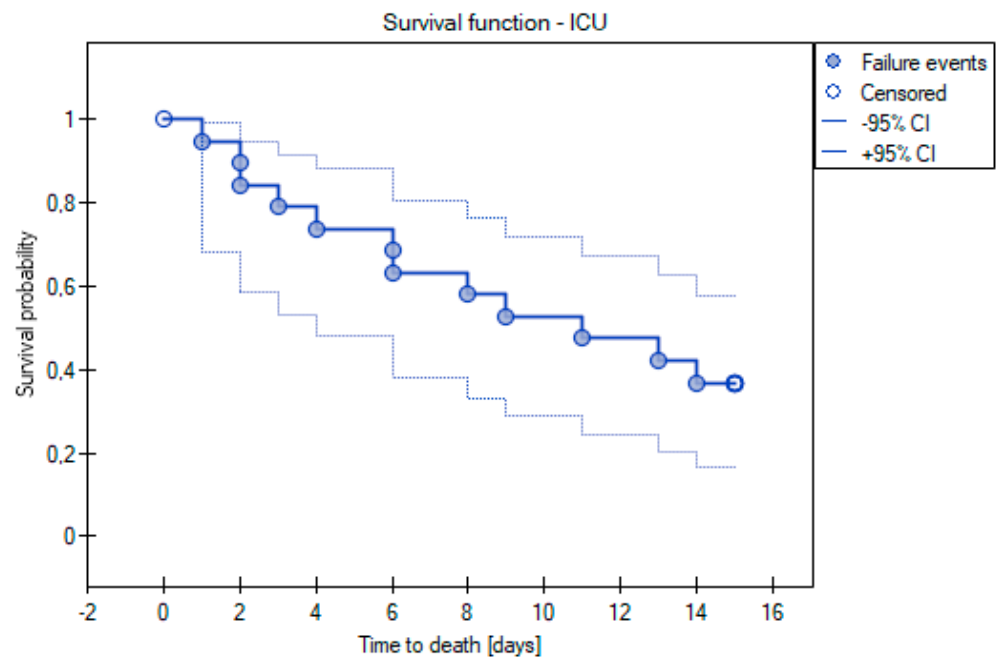


Figure 4. Kaplan–Meier curve for survival time analysis for ICU cases.

By using the Kaplan–Meier curve, we also decided to examine the median survival time of the patients using the PCR test. Kaplan–Meier curves were applied to examine the median survival time of the patients using the PCR test. Only the patients who died within 2 months of the PCR test were included in this analysis. Based on this curve, a median value of 10 days and a mean value of 14 days were determined (Figure 5).

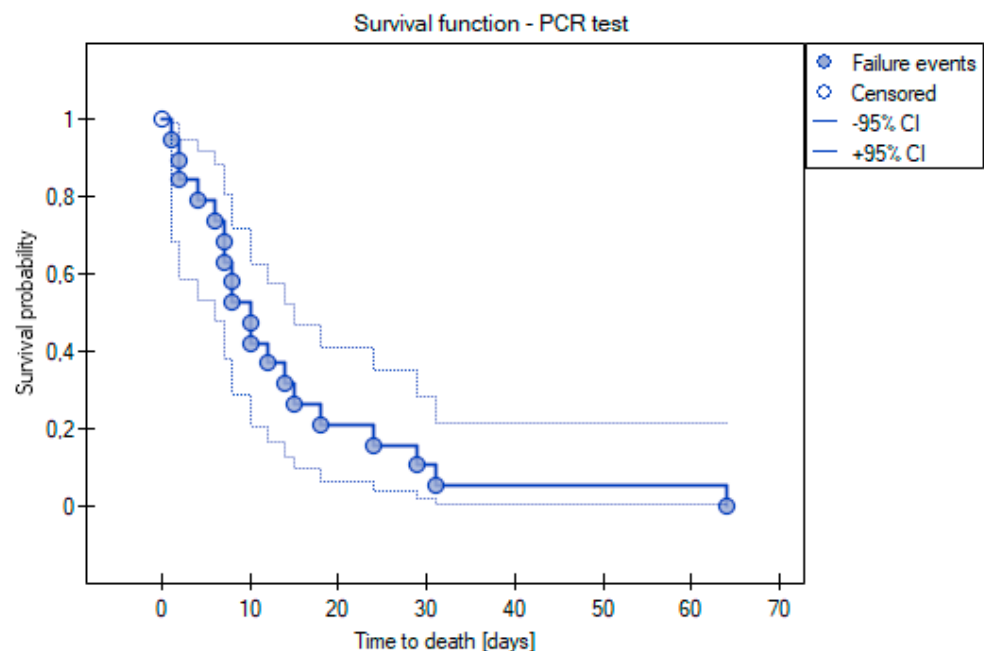


Figure 5. Kaplan–Meier curve for survival time analysis for PCR test cases.

3.2. Inter-Observer Agreement

For each scoring system, 100 observations were submitted.

There was statistically significant inter-observer agreement between the two observers in assessing qualitative lung involvement using the m-TSS method. The inter-observer concordance oscillated at the level of perfect agreement (kappa value = 1) for four categories, i.e., GGO, crazy paving, consolidations, and normal lungs (Table 2).

Table 2. Test of significance: Fleiss’ kappa for qualitative lung assessment via m-TSS performed by the two observers.

Category	Fleiss’ Kappa (κ)	95% CI	SE	p Value
Overall	1	0.796–1	0.104	<0.001
Nomal lungs	1	0.723–1	0.141	<0.001
Ground glass opacities	1	0.723–1	0.141	<0.001
Crazy paving	1	0.723–1	0.141	<0.001
Consolidations	1	0.723–1	0.141	<0.001

κ —kappa value, CI—Confidence interval, SE—standard error.

As part of the inter-observer agreement analysis, an error plot was determined for the two observers (Figure 6), which graphically shows the same values shown in Table 2.

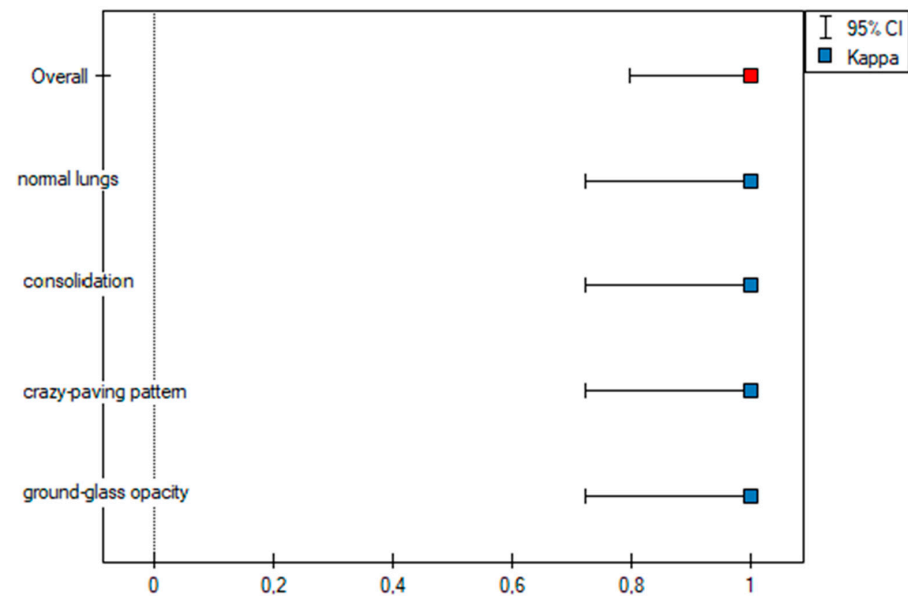


Figure 6. Error plot for the m-TSS method.

In addition, an agreement plot (Figure 7) was determined, which shows how the patient from the first to the fiftieth was rated by the two observers, where their agreement was 100%, which is why each column in the plot is the same color.

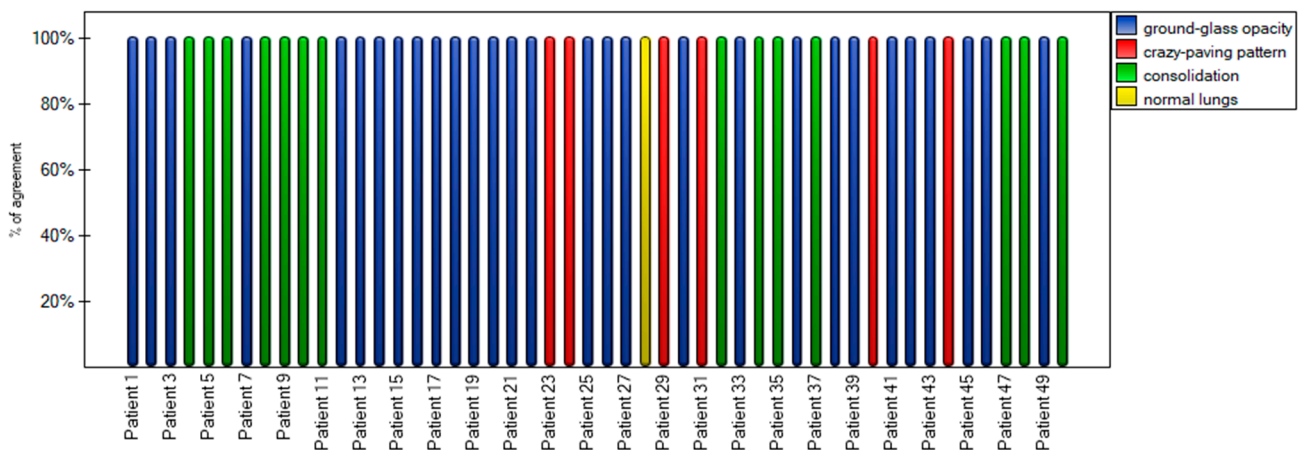


Figure 7. Agreement stacked column plot for the m-TSS method.

Based on the ICC values, excellent inter-observer reliability was found among the three methods, where ICC > 0.9 (Table 3).

Table 3. Inter-rater reliability for quantitative scoring systems.

Scoring System	ICC	95% CI	p Value
Chest CT severity score	0.994	0.99–0.997	<0.001
Chest CT score	0.994	0.99–0.997	<0.001
Total severity score	0.992	0.987–0.996	<0.001

ICC—Intraclass correlation coefficient, CI—Confidence interval.

To graphically represent the inter-observer agreement between the two observers in the quantitative methods shown in Table 3, multiple dot graphs were created (Figures 8–10).

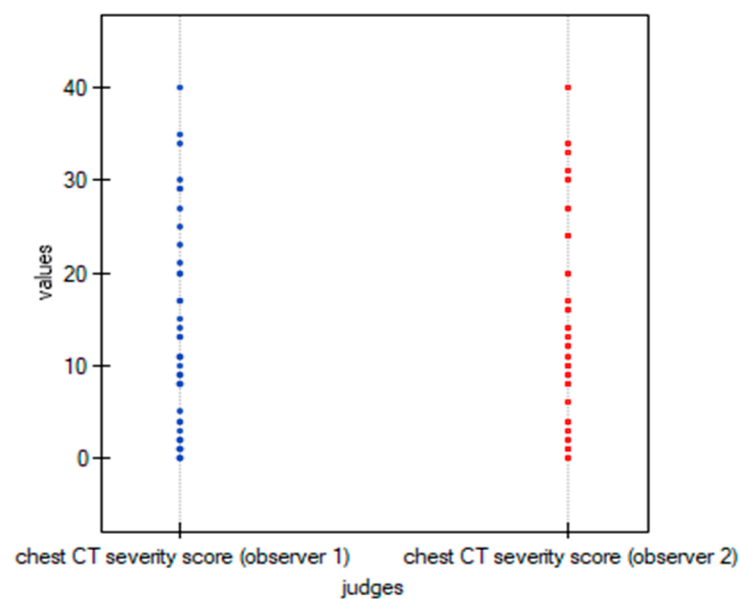


Figure 8. Multiple dot graphs for inter-observer agreement for Chest CT-SS.

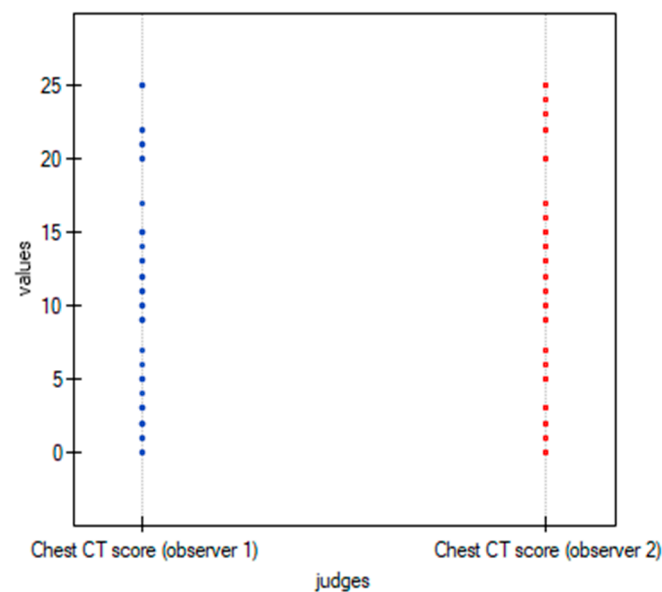


Figure 9. Multiple dot graphs for inter-observer agreement for Chest CT-S.

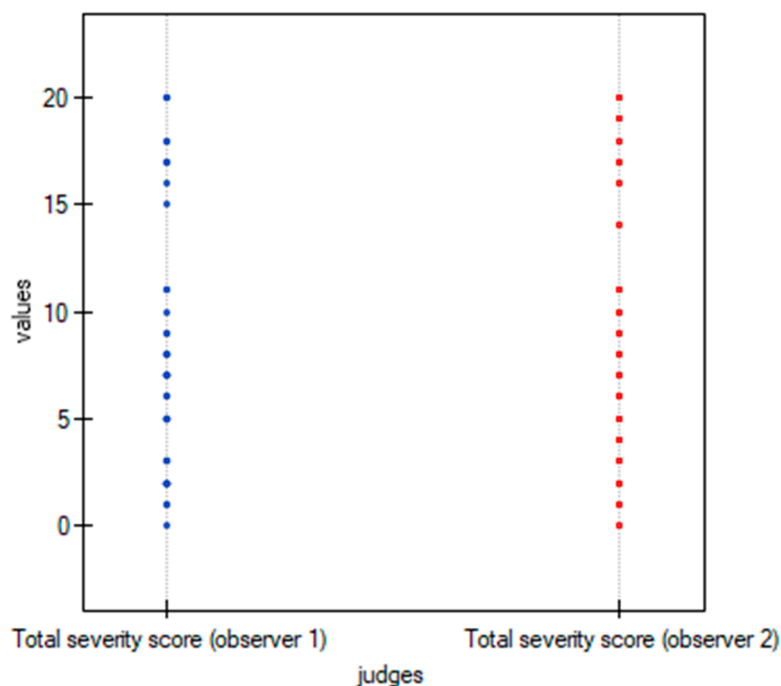


Figure 10. Multiple dot graphs for inter-observer agreement for TSS.

3.3. Severity Scoring Systems

The ROC curve was used to determine which of the three quantitative methods would best assign patients to the two groups according to the severity of COVID-19 disease. The two groups were determined: non-severe and severe, as described in the Materials and Methods section.

Table 4 shows all the values important for the comparison and analysis regarding the ROC curve for the three quantitative methods, i.e., range of values, sensitivity, specificity, cutoff value, AUC value, and *p*-value.

Table 4. Sensitivity and specificity of the quantitative methods for identifying the probability of COVID-19 adverse outcomes in the studied patients (severe and non-severe). The results were obtained based on the analysis of the ROC curves.

Scoring System	Range	Sensitivity	Specificity	Cutoff Value	AUC	<i>p</i> Value
Chest CT severity score	0–40	72.7%	98.2%	≥14	0.902	<0.001
Chest CT score	0–25	75%	100%	≥14	0.899	<0.001
Total severity score	0–20	65.9%	94.6%	≥9	0.881	<0.001

By using ROC curves (Figures 11–13), a selection of the optimal cutoff value was made, i.e., a certain value of the diagnostic variable that best divides the study population into two groups: severe and non-severe COVID-19. Comparing the results obtained for the ROC curves, it can be seen that for the two quantitative methods (Chest CT-SS and Chest CT-S), the cutoff value is the same at ≥ 14 , and for the TSS method it is ≥ 9 .

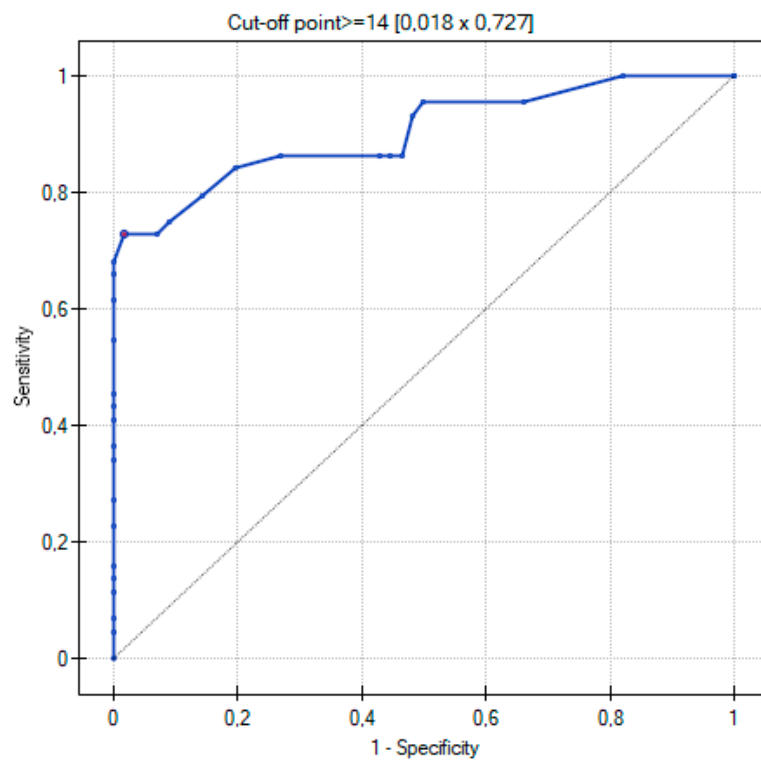


Figure 11. ROC curve for the diagnostic performance of Chest CT-SS in detection of severe COVID-19 cases.

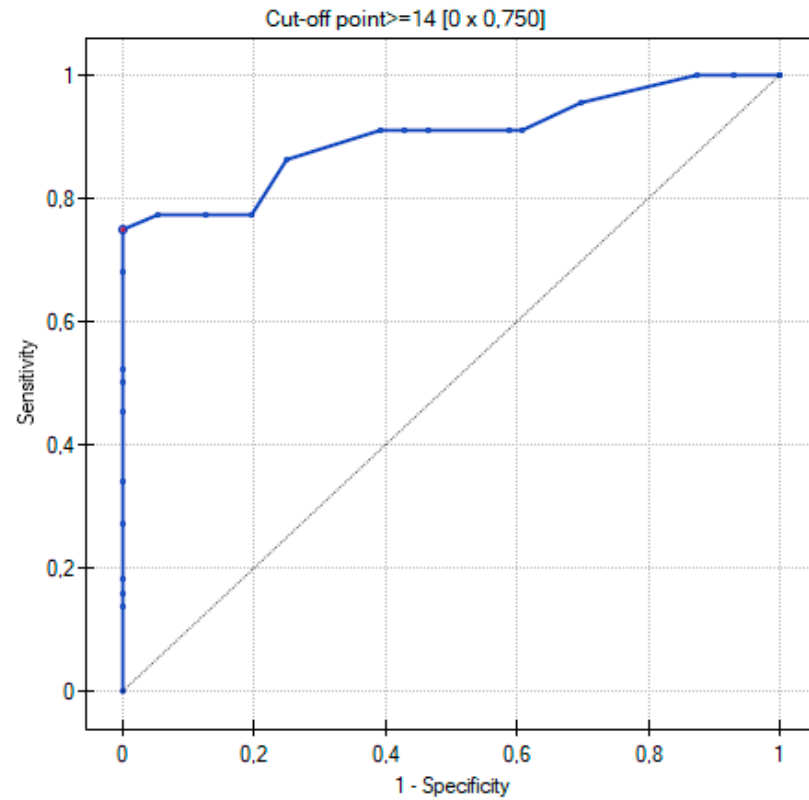


Figure 12. ROC curve for the diagnostic performance of Chest CT-S in detection of severe COVID-19 cases.

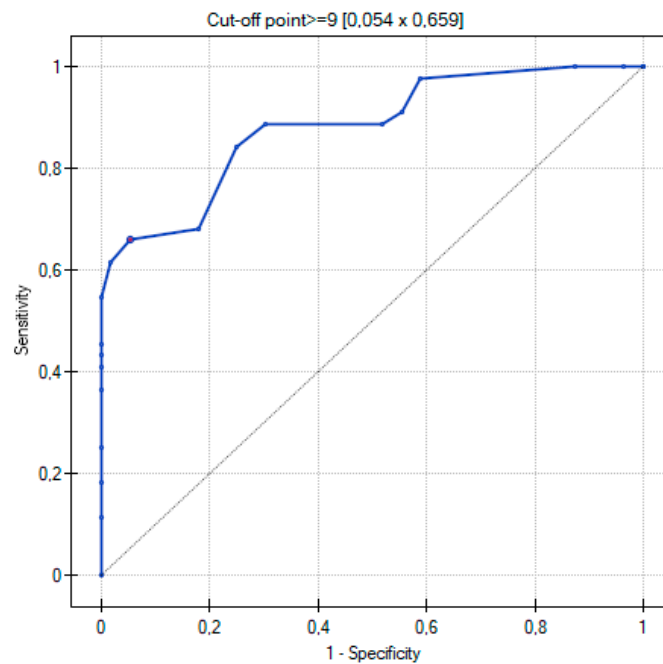


Figure 13. ROC curve for the diagnostic performance of TSS in detection of severe COVID-19 cases.

The three ROC curves revealed excellent and very good diagnostic accuracy for the three scoring systems. The AUC values were excellent (0.902), very good (0.899), and very good (0.881) for the CT-SS, CT-S, and TSS scoring systems, respectively. Sensitivity showed high levels at 72.7%, 75%, and 65.9% respectively, and specificity was recorded at 98.2%, 100%, and 94.6% for the CT-SS, CT-S, and TSS scoring systems, respectively.

In order to be able to compare the mean values obtained when the two observers performed the three scoring methods for severe and non-severe COVID-19, a *t*-test for independent samples was performed. With the obtained *p*-value < 0.05, box plots (Figures 14–16) were created to graphically visualize the obtained results. The same results are also presented in Table 5.

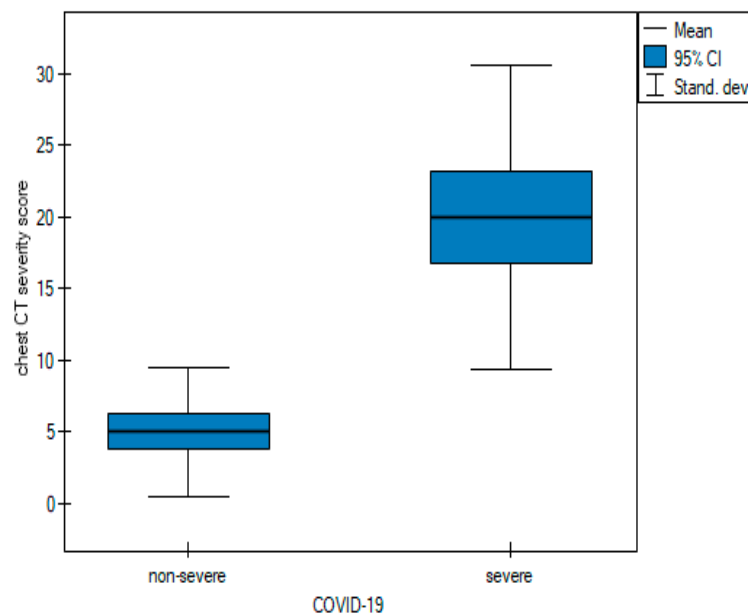


Figure 14. Box plot of comparisons of mean chest CT-SS for non-severe and severe COVID-19 patients.

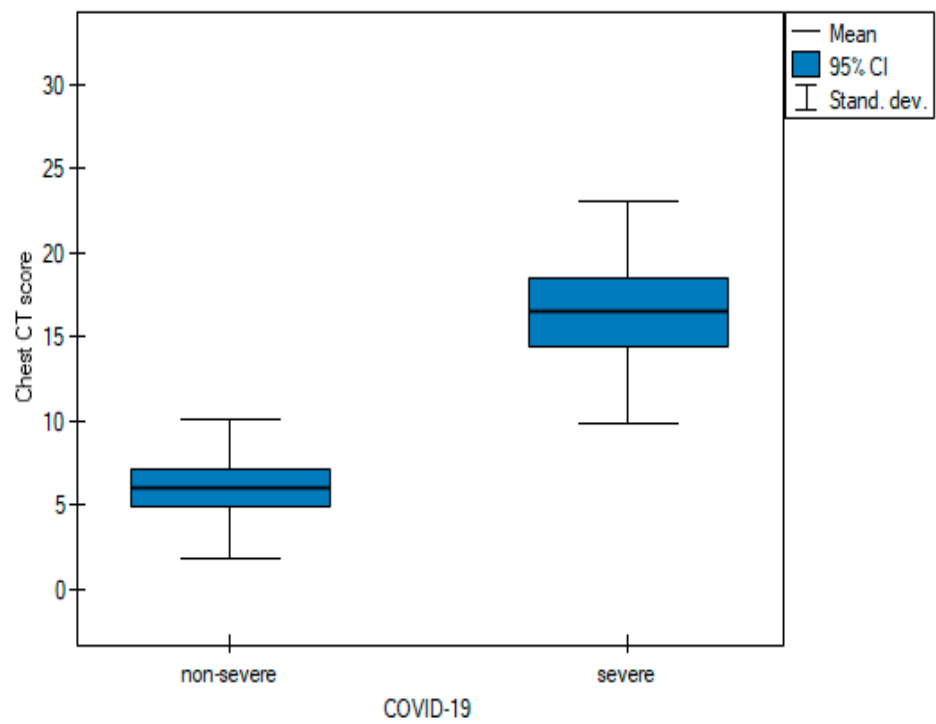


Figure 15. Box plot of comparisons of mean Chest CT-S for non-severe and severe COVID-19 patients.

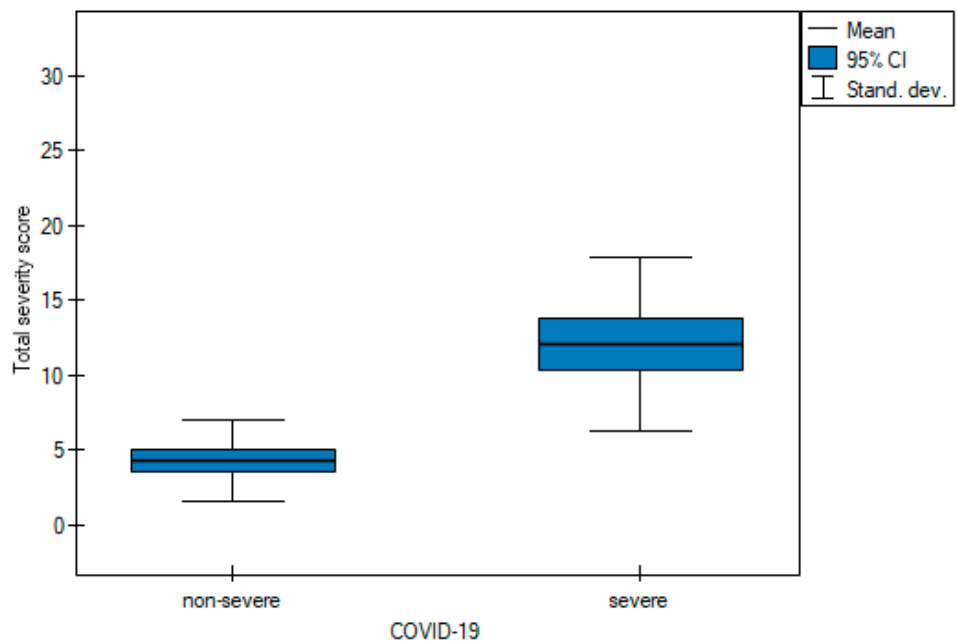


Figure 16. Box plot of comparisons of mean TSS for non-severe and severe COVID-19 patients.

Table 5. The mean values of the three quantitative methods obtained with the tests conducted by the two observers.

Characteristic	Total	Non-Severe COVID-19	Severe COVID-19	p Value
Chest CT severity score	11.6 ± 10.78	5.02 ± 4.53	19.98 ± 10.64	<0.05 ^t
Chest CT score	10.62 ± 7.49	6 ± 4.15	16.5 ± 6.63	<0.05 ^t
Total severity score	7.74 ± 5.81	4.3 ± 2.68	12.11 ± 5.8	<0.05 ^t

Test of significance: ^t independent-samples *t* test.

A time consumption analysis (Table 6) of each of the three methods was used to decide which method should be chosen first.

Table 6. Comparisons of examination time in seconds of each of the three quantitative scoring systems.

Interpretation Time [seconds]				
Statistic	Chest CT Severity Score	Chest CT Score	Total Severity Score	<i>p</i> Value
Average	35.32	35.62	29.86	0.002
Median	30	35	30	0.002
Q ₁ –Q ₃	22.75–40	28–45	22–38	
Range	10–78	10–68	10–55	
Pairwise comparisons	B	B	A	

p value: Friedman test, Q₁–Q₃: the first quartile–the third quartile. Pairwise comparisons (Test POST-HOC Conover-Iman): similar letters = insignificant difference, different letters = significant difference.

3.4. Time Consumption

Table 6 shows the results obtained during the analysis, which was based on the average and median time of the study.

The results arranged from the longest to the shortest according to the average were as follows: Chest CT score_time > Chest CT severity score_time > Total severity score_time. The results arranged according to the median value: Chest CT score > Chest CT severity score = Total severity score. In addition, pairwise comparisons showed a statistically significant difference between all pairs except Chest CT severity score_time and Chest CT severity score_time, which ended up in the same group (B). Figure 17 shows box plots comparing the mean interpretation time of the three quantitative scoring systems, whereas Figure 18 compares the median time.

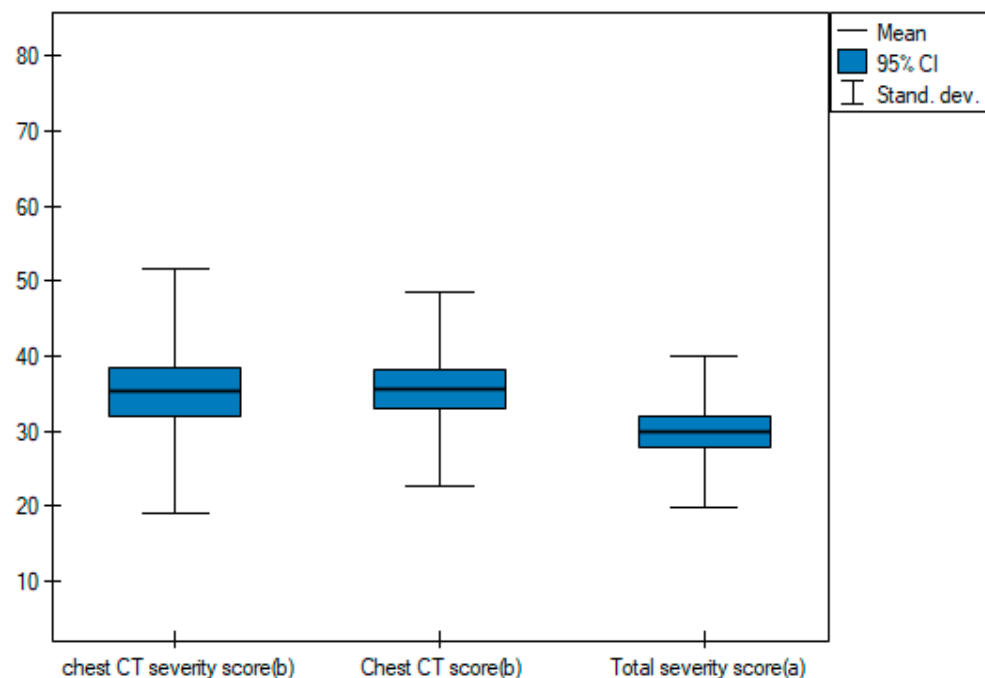


Figure 17. Box plot of comparisons of the mean interpretation times of the three quantitative scoring systems.

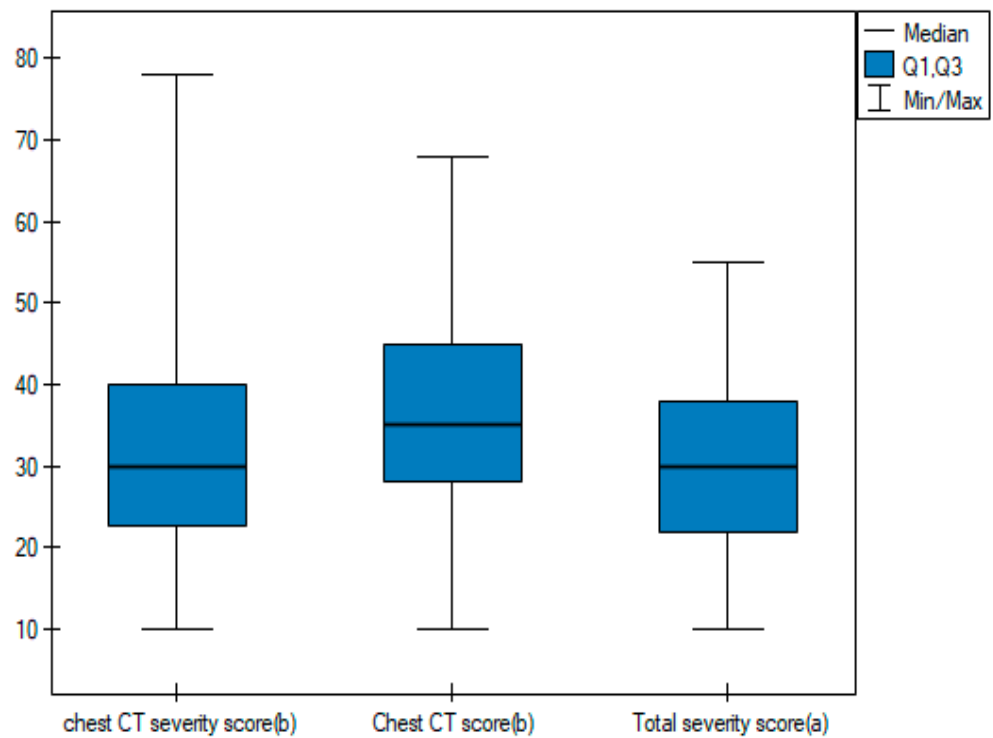


Figure 18. Box plot of comparisons of the median interpretation times of the three quantitative scoring systems.

Figures 19–21 illustrate typical manifestations of COVID-19 pneumonia in analyzed patients.

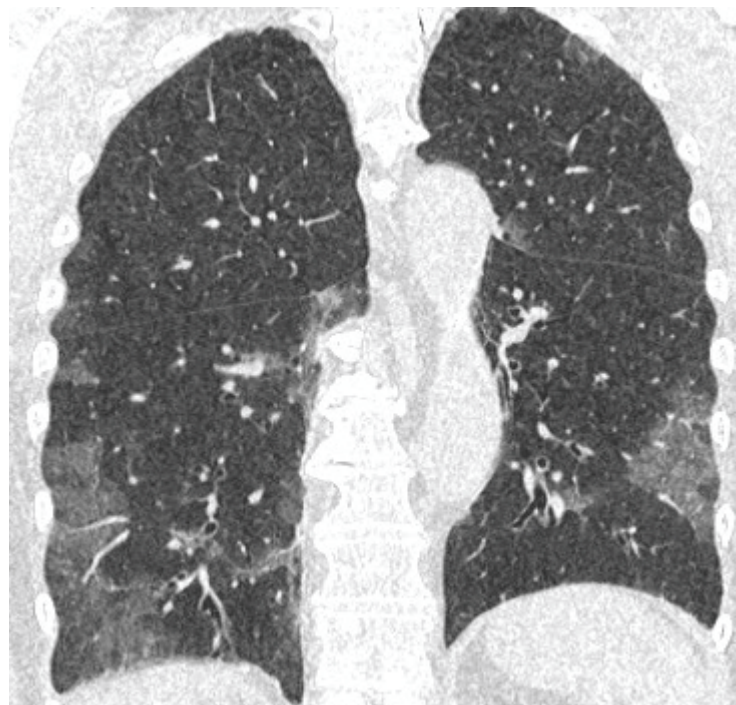


Figure 19. Coronal reconstruction of non-enhanced CT reveals bilateral ground glass opacities in peripheral distribution in lower lobes.

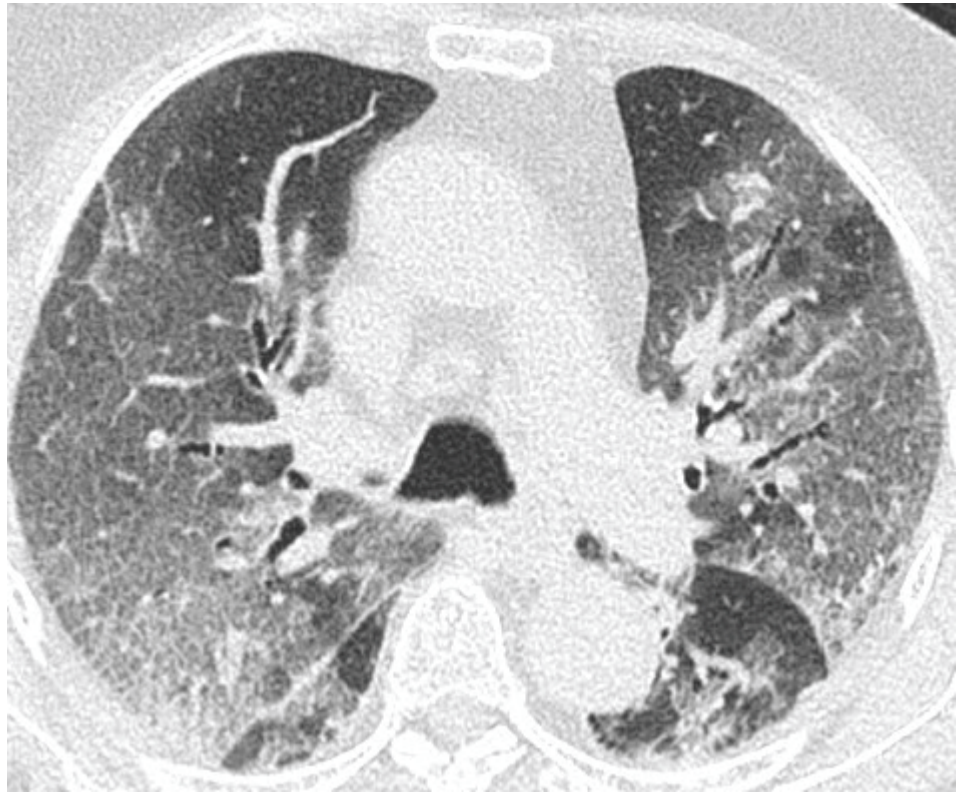


Figure 20. Axial plane of non-enhanced CT reveals bilateral ground glass opacities with superimposition of interlobular septal thickening, giving the crazy paving appearance, distributed symmetrically in both lungs.

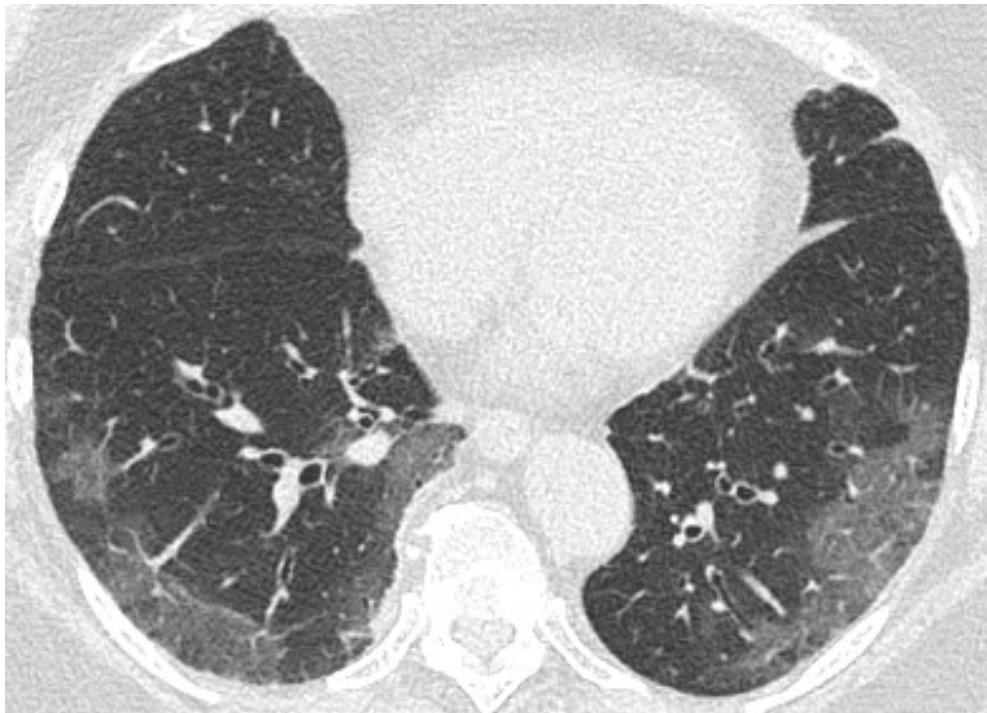


Figure 21. Axial plane of non-enhanced CT reveals ground glass opacities in subpleural distribution in lower lobes.

4. Discussion

The role of chest CT in the pandemic era of COVID-19 has been established as an auxiliary method for assessing the severity and extent of lung involvement in a selected group of patients. The additional use of a scoring system for evaluation allows for more objective and standardized results and may also be useful as a short-term prognostic factor [14]. There have been several studies evaluating scoring systems in COVID-19 pneumonia in different groups of patients [5,6,11,14].

According to Tharwat et. al, TSS may have prognostic value in patients with acute renal failure or chronic kidney disease. Higher TSS scores and a predominant pattern of pulmonary consolidations were more common in patients in severe clinical condition ($p < 0.05$) [15]. These findings are consistent with our study, with significantly higher values in TSS in the severe COVID-19 group.

The m-TSS method was a qualitative method to distinguish the main radiological patterns present in COVID-19 pneumonia. Category A for the GGO pattern was the most common observation in our study and predominated in 58% of patients (Figures 19–21). GGO is not a feature specific to COVID-19 pneumonia and can occur in several conditions. In hematologic patients, the differential diagnosis of GGO should include other viral infections such as cytomegalovirus, alveolar hemorrhage, drug toxicity and organizing pneumonia [16]. Our study supports the results of the meta-analysis by Zheng et al., where GGO, vascular enlargement, alveolar septal thickening, and subpleural lines were the most common findings in both normal and severe patients [17]. Furthermore, our results presented excellent inter-observer agreement in qualitative analysis in the mTSS method ($p < 0.001$).

In a study by Yang et al., the threshold for identifying patients with severe infection in CT-SS was 19.5 with 83.3% sensitivity and 94% specificity [6]. In our study, the results for CT-SS were 14, 72.7%, and 98.2% respectively, confirming excellent diagnostic performance of the method. Additionally, the AUC values of analyzed semi-quantitative methods ranged between 0.881 and 0.902, suggesting that all of them can be considered as excellent tools for discriminating patients with severe and non-severe COVID-19 disease. In all semi quantitative methods, patients in the severe COVID-19 group received higher scores than non-severe patients ($p < 0.005$). In the study by Li et al., patients with severe clinical conditions had significantly higher scores in the CT-S. Moreover, sensitivity, specificity, and cut-off values for identifying severe cases were 80%, 82.8%, and 7, respectively, whereas the AUC value for the ROC curve was 0.87 [18]. In our study, the values were as follows: 75%, 100%, and 14. The AUC value in the ROC curve was 0.899. Our results are similar, confirming very good diagnostic performance of this scale, even though the analyzed group was smaller. Another important point of analysis is the repeatability and feasibility of the method. In our study, we achieved excellent inter-observer reliability in all three semi quantitative methods with the ICC values > 0.9 .

During the pandemic era, several CT scales were invented to assess lung involvement in COVID-19 infection. In a recent study published in May 2022, Dilek et al. proposed an Early Decision Severity Score. The system includes evaluation of patients based on visual CT scorings, intubation necessities, and mortality rates. The combination of radiological and clinical factors had the additional benefit of assessing patient's prognosis [19]. In our study, several clinical parameters were analyzed.

Respiratory failure is one of the leading causes of COVID-19 mortality [20]. Patients with low oxygen saturation have a higher risk of death, which supports results that link hypoxemia to mortality [21]. In the analyzed group of patients, oxygen saturation varied significantly with the average values in non-severe and severe patients reaching 94% and 82%, respectively (Table 1 and Figure 1).

One of the laboratory parameters associated with mortality in patients with COVID-19 is D-dimer. D-dimer levels are elevated in patients of all age groups with COVID-19. In patients with COVID-19, the existence of a concomitant disease such as diabetes, cancer, stroke, and physiological conditions such as pregnancy may contribute to elevated D-dimer

levels. On the other hand, the correlation between high D-dimer levels and survival rates underscores the importance of detecting D-dimer levels in patients with COVID-19 [22]. In our study, not surprisingly, D-dimer values were significantly higher in the severe COVID-19 group with the mean value of 3562 µg/L (Table 1 and Figure 2).

CRP levels were found to be significantly elevated in the initial phases of the infection in patients with severe COVID-19 also prior to indications of critical findings with CT. Importantly, CRP has been associated with disease progression and is an early predictor for severe COVID-19 [23]. In our study, patients in the severe COVID-19 group had significantly higher CRP levels in contrast to the non-severe COVID-19 group (Table 1 and Figure 3).

Compared to the general patient population, those with comorbidities have a poor prognosis and high mortality resulting from COVID-19 [24,25]. In hematological patients, advanced disease, older age, type of malignancy, and several laboratory parameters, such as high CRP, lymphopenia, and neutropenia have been correlated with COVID-19 mortality [26–28]. Chronic diseases such as diabetes or arterial hypertension are associated with a high mortality rate [29]. Our study supports those findings, with significantly higher mortality in patients with comorbidities, particularly in those with diabetes.

To the best of our knowledge, this is the first study analyzing diagnostic accuracy of chest CT scales in COVID-19 in hematology patients. A recent study by Elmokadem et al. analyzed the diagnostic performance of five different CT chest severity scoring systems in ordinary patients and found that Chest CT-SS had the highest specificity and utilized the least amount of time when compared to other scoring systems [30]. The results were concordant with our study, with 98.2% specificity for Chest CT-SS and excellent AUC values of 0.902 ($p < 0.001$). Interestingly, the median interpretation time in Chest CT-SS and TSS was equal, while in Chest CT-S, it was slightly longer ($p = 0.002$).

Our study has a few limitations. Firstly, as it was a single-center study, the group was relatively small and thus further studies with more patients are warranted. To confirm our results, larger datasets are needed. Secondly, patients suffered from different hematological malignancies at different stages of diseases. Taking this fact into account, some laboratory parameters could not be reliably interpreted and collected for analysis due to different clinical stages of hematological disease. Furthermore, due to the retrospective nature of the study, some clinical data were incomplete and thus clinical analysis was very limited to selected laboratory parameters; however, we managed to collect one significant to assess prognosis and risk factors such as D-dimer and CRP.

5. Conclusions

The analyzed chest CT scores presented excellent inter-observer agreement. Chest CT score and chest CT severity score presented very high sensitivity and specificity in terms of diagnostic accuracy. The highest AUC values and shortest median times of analysis in chest CT severity scores suggests this method as the most comprehensive for semi-quantitative assessment of chest CT in hematological patients with COVID-19.

Author Contributions: Conceptualization, M.H.-R. and L.G.-D.; methodology, M.H.-R. and L.G.-D.; software, M.H.-R. and L.G.-D.; validation, M.H.-R., L.G.-D., J.D.-S. and M.J.; formal analysis, M.H.-R., L.G.-D. and J.D.-S.; investigation, M.H.-R., L.G.-D. and J.D.-S.; resources M.H.-R. and L.G.-D.; data curation, M.H.-R. and L.G.-D.; writing—original draft preparation, M.H.-R. and L.G.-D.; writing—review and editing, M.H.-R., L.G.-D. and M.J.; visualization, M.H.-R. and L.G.-D.; supervision, L.G.-D. and M.J.; project administration, M.H.-R. and L.G.-D. 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 in accordance with the Declaration of Helsinki and approved by the Institutional Review Board of Medical University of Warsaw AKBE/297/2022 date 12 December 2022.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available in this article.

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

References

1. Leung, C. Risk factors for predicting mortality in elderly patients with COVID-19: A review of clinical data in China. *Mech. Ageing Dev.* **2020**, *188*, 111255. [[CrossRef](#)]
2. Tagliamento, M.; Agostinetto, E.; Bruzzzone, M.; Ceppi, M.; Saini, K.S.; de Azambuja, E.; Punie, K.; Westphalen, C.B.; Morgan, G.; Pronzato, P.; et al. Mortality in adult patients with solid or hematological malignancies and SARS-CoV-2 infection with a specific focus on lung and breast cancers: A systematic review and meta-analysis. *Crit. Rev. Oncol. Hematol.* **2021**, *163*, 103365. [[CrossRef](#)]
3. Mossa-Basha, M.; Meltzer, C.C.; Kim, D.C.; Tuite, M.J.; Kolli, K.P.; Tan, B.S. Radiology Department preparedness for COVID-19: Radiology Scientific Expert Panel. *Radiology* **2020**, *296*, E106–E112. [[CrossRef](#)]
4. Cieszanowski, A.; Czekańska, E.; Giżycka, B.; Gruszczyńska, K.; Oronowicz-Jaśkowiak, A.S.; Podgórska, J.; Serafin, Z.; Szurowska, E.; Walecki, J. Indications for imaging studies in SARS-CoV-2 infected patients—Recommendations of the Polish Medical Society of Radiology. *Pol. J. Radiol.* **2022**, *87*, e63–e68. [[CrossRef](#)]
5. Ai, T.; Yang, Z.; Hou, H.; Zhan, C.; Chen, C.; Lv, W.; Tao, Q.; Sun, Z.; Xia, L. Correlation of chest CT and RT-PCR testing in coronavirus disease 2019 (COVID-19) in China: A report of 1014 cases. *Radiology* **2020**, *296*, 200642. [[CrossRef](#)]
6. 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. [[CrossRef](#)]
7. Eshet, Y.; Avigdor, A.; Kedmi, M.; Tau, N. Imaging of Hematological Patients in the Era of COVID-19. *Acta Haematol.* **2022**, *145*, 267–274. [[CrossRef](#)]
8. Flisiak, R.; Horban, A.; Jaroszewicz, J.; Kozielowicz, D.; Mastalerz-Migas, A.; Owczuk, R.; Parczewski, M.; Pawłowska, M.; Piekarska, A.; Simon, K.; et al. Management of SARS-CoV-2 infection: Recommendations of the Polish Association of Epidemiologists and Infectiologists as of April 26, 2021. *Pol. Arch. Intern. Med.* **2021**, *131*, 487–496. [[CrossRef](#)]
9. Austin, J.H.; Muller, N.L.; Friedman, P.J.; Hansell, D.M.; Naidich, D.P.; Remy-Jardin, M.; Webb, W.R.; Zerhouni, E.A. Glossary of terms for CT of the lungs: Recommendations of the Nomenclature Committee of the Fleischner Society. *Radiology* **1996**, *200*, 327–331. [[CrossRef](#)]
10. Pan, F.; Ye, T.; Sun, P.; Gui, S.; Liang, B.; Li, L.; Zheng, D.; Wang, J.; Hesketh, R.L.; Yang, L.; et al. Time course of lung changes on chest CT during recovery from 2019 novel coronavirus (COVID-19) pneumonia. *Radiology* **2020**, *295*, 715–721. [[CrossRef](#)]
11. Francone, M.; Iafrate, F.; Masci, G.M.; Coco, S.; Cilia, F.; Manganaro, L.; Panebianco, V.; Andreoli, C.; Colaiacomo, M.C.; Zingaropoli, M.A.; et al. Chest CT score in COVID-19 patients: Correlation with disease severity and short-term prognosis. *Eur. Radiol.* **2020**, *30*, 6808–6817. [[CrossRef](#)]
12. Li, K.; Fang, Y.; Li, W.; Pan, C.; Qin, P.; Zhong, Y.; Liu, X.; Huang, M.; Liao, Y.; Li, S. CT image visual quantitative evaluation and clinical classification of coronavirus disease (COVID-19). *Eur. Radiol.* **2020**, *30*, 4407–4416. [[CrossRef](#)]
13. Wasilewski, P.G.; Mruk, B.; Mazur, S.; Póttorak-Szymczak, G.; Sklinda, K.; Walecki, J. COVID-19 severity scoring systems in radiological imaging—A review. *Pol. J. Radiol.* **2020**, *85*, e361. [[CrossRef](#)] [[PubMed](#)]
14. Aziz-Ahari, A.; Keyhanian, M.; Mamishi, S.; Mahmoudi, S.; Bastani, E.E.; Asadi, F.; Khaleghi, M. Chest CT severity score: Assessment of COVID-19 severity and short-term prognosis in hospitalized Iranian patients. *Wien. Med. Wochenschr.* **2022**, *172*, 77–83. [[CrossRef](#)] [[PubMed](#)]
15. Tharwat, S.; Saleh, G.A.; Saleh, M.; Mounir, A.M.; Abdelzaher, D.G.; Salah, A.M.; Nassar, M.K. Chest CT Total Severity Score on Admission to Predict In-Hospital Mortality in COVID-19 Patients with Acute and Chronic Renal Impairment. *Diagnostics* **2022**, *12*, 1529. [[CrossRef](#)]
16. Cozzi, D.; Cavigli, E.; Moroni, C.; Smorchkova, O.; Zantonelli, G.; Pradella, S.; Miele, V. Ground-glass opacity (GGO): A review of the differential diagnosis in the era of COVID-19. *Jpn. J. Radiol.* **2021**, *39*, 721–732. [[CrossRef](#)] [[PubMed](#)]
17. Zheng, Y.; Wang, L.; Ben, S. Meta-analysis of chest CT features of patients with COVID-19 pneumonia. *J. Med. Virol.* **2021**, *93*, 241–249. [[CrossRef](#)]
18. Li, K.; Wu, J.; Wu, F.; Guo, D.; Chen, L.; Fang, Z.; Li, C. The Clinical and Chest CT Features Associated with Severe and Critical COVID-19 Pneumonia. *Investig. Radiol.* **2020**, *55*, 327–331. [[CrossRef](#)] [[PubMed](#)]
19. Dilek, O.; Demirel, E.; Akkaya, H.; Belibagli, M.C.; Soker, G.; Gulek, B. Different chest CT scoring systems in patients with COVID-19: Could baseline CT be a helpful tool in predicting survival in patients with matched ages and co-morbid conditions? *Acta Radiol.* **2022**, *63*, 615–622. [[CrossRef](#)]
20. Lai, A.G.; Pasea, L.; Banerjee, A.; Hall, G.; Denaxas, S.; Chang, W.H.; Katsoulis, M.; Williams, B.; Pillay, D.; Noursadeghi, M.; et al. Estimated Impact of the COVID-19 Pandemic on Cancer Services and Excess 1-Year Mortality in People with Cancer and Multimorbidity: Near Real-Time Data on Cancer Care, Cancer Deaths and a Population-Based Cohort Study. *BMJ Open* **2020**, *10*, e043828. [[CrossRef](#)]
21. Ferrer, R. COVID-19 Pandemic: The Greatest Challenge in the History of Critical Care. *Med. Intensiva* **2020**, *44*, 323–324. [[CrossRef](#)] [[PubMed](#)]
22. Rostami, M.; Mansouritorghabeh, H. D-dimer level in COVID-19 infection: A systematic review. *Expert Rev. Hematol.* **2020**, *11*, 1265–1275. [[CrossRef](#)]

23. Tan, C.; Huang, Y.; Shi, F.; Tan, K.; Ma, Q.; Chen, Y.; Jiang, X.; Li, X. C-reactive protein correlates with CT findings and predicts severe COVID-19 early. *J. Med. Virol.* **2020**, *92*, 856–862. [[CrossRef](#)]
24. Wang, C.; Horby, P.W.; Hayden, F.G.; Gao, G.F. A novel coronavirus outbreak of global health concern. *Lancet* **2020**, *395*, 470–473. [[CrossRef](#)] [[PubMed](#)]
25. Bernheim, A.; Mei, X.; Huang, M.; Yang, Y.; Fayad, Z.A.; Zhang, N.; Diao, K.; Lin, B.; Zhu, X.; Li, K.; et al. Chest CT findings in coronavirus disease-19 (COVID-19): Relationship to duration of infection. *Radiology* **2020**, *295*, 200463. [[CrossRef](#)] [[PubMed](#)]
26. García-Suárez, J.; de la Cruz, J.; Cedillo, A.; Llamas, P.; Duarte, R.; Jiménez-Yuste, V.; Hernández-Rivas, J.Á.; Gil-Manso, R.; Kwon, M.; Sánchez-Godoy, P.; et al. Impact of hematologic malignancy and type of cancer therapy on COVID-19 severity and mortality: Lessons from a large population-based registry study. *J. Hematol. Oncol.* **2020**, *13*, 133. [[CrossRef](#)]
27. Piñana, J.L.; Martino, R.; García-García, I.; Parody, R.; Morales, M.D.; Benzo, G.; Gómez-Catalan, I.; Coll, R.; De La Fuente, I.; Luna, A.; et al. Risk factors and outcome of COVID-19 in patients with hematological malignancies. *Exp. Hematol. Oncol.* **2020**, *9*, 21. [[CrossRef](#)]
28. Passamonti, F.; Cattaneo, C.; Arcaini, L.; Bruna, R.; Cavo, M.; Merli, F.; Angelucci, E.; Krampera, M.; Cairoli, R.; Della Porta, M.G.; et al. Clinical characteristics and risk factors associated with COVID-19 severity in patients with hematological malignancies in Italy: A retrospective, multicentre, cohort study. *Lancet Haematol.* **2020**, *7*, e737–e745. [[CrossRef](#)]
29. Di Lorenzo, G.; Buonerba, L.; Ingenito, C.; Crocetto, F.; Buonerba, C.; Libroia, A.; Sciarra, A.; Ragone, G.; Sanseverino, R.; Iaccarino, S.; et al. Clinical Characteristics of Metastatic Prostate Cancer Patients Infected with COVID-19 in South Italy. *Oncology* **2020**, *98*, 743–747. [[CrossRef](#)]
30. Elmokadem, A.H.; Mounir, A.M.; Ramadan, Z.A.; Elsedeyq, M.; Saleh, G.A. Comparison of Chest CT Severity Scoring Systems for COVID-19. *Eur. Radiol.* **2022**, *32*, 3501–3512. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.