1	Quantitative analysis of chest computed tomography of COVID-19 pneumonia using a software
2	widely used in Japan
3	
4	Minako Suzuki ^{1*} , Yoshimi Fujii ^{2¶} , Yurie Nishimura ^{2¶} , Kazuma Yasui ^{2¶} , Hidefumi Fujisawa ^{1¶}
5	
6	¹ Department of Radiology, Showa University Northern Yokohama Hospital, Yokohama, Kanagawa,
7	Japan
8	² Department of Radiology, Fujisawa City Hospital, Fujisawa, Kanagawa, Japan
9	
10	* Corresponding author:
11	E-mail: minacom69@gmail.com (MS)
12	
13	Short title: Quantitative analysis of chest CT of COVID-19 pneumonia
14	
15	[¶] These authors contributed equally to this work.

17 Abstract

18 This study aimed to determine the optimal conditions to measure the percentage of area 19 considered as pneumonia (pneumonia volume ratio, PVR) and the computed tomography 20 (CT) score due to coronavirus disease 2019 (COVID-19) using the Ziostation2 image 21 analysis software (Z2; Ziosoft, Tokyo, Japan), which is popular in Japan, and to evaluate its 22 usefulness in assessing the clinical severity. We included 53 patients (41 men and 12 women, 23 mean age: 61.3 years) diagnosed with COVID-19 using the polymerase chain reaction who 24 had undergone chest CT and were hospitalized between January 2020 and January 2021. 25 Based on the COVID-19 infection severity, the patients were classified as mild (n=38) or 26 severe (n=15). For 10 randomly selected samples, the PVR and CT scores by Z2 under 27 different conditions and the visual simple PVR and CT scores were compared, and the 28 conditions with the highest statistical agreement were determined. The usefulness of the clinical severity assessment based on PVR and CT scores using Z2 under the determined 2930 conditions was statistically evaluated. The best agreement with the visual measurement was 31 achieved by the Z2 measurement condition of ≥ -600 HU. The areas under the receiver 32 operating characteristic curves, the Youden index, and the sensitivity, specificity, and p-values 33 of PVR and CT scores by Z2 were as follows: PVR; 0.881, 18.69, 66.7, 94.7, and <0.001, CT 34 score; 0.77, 7.5, 40, 74, and 0.002, respectively. We determined the optimal condition for 35 assessing the PVR of COVID-19 pneumonia using Z2 and demonstrated that the AUC of 36 PVR was higher than that of the CT score in the assessment of clinical severity. The 37 introduction of new technologies is time-consuming and expensive; our method has high 38 clinical utility and can be promptly used in any facility where Z2 has been introduced.

39

40 Introduction

41	The coronavirus disease 2019 (COVID-19) pandemic caused by the novel coronavirus severe
42	acute respiratory syndrome coronavirus 2 (SARS-CoV-2) was first identified in Wuhan,
43	China, and reported in December 2019 [1]. The pandemic prevails with an increasing number
44	of infections and deaths. In Japan, the first COVID-19 case was reported in January 2020 [2],
45	and by April 2023, over 33 million people had been infected, and more than 74000 people
46	had died [3]. After the Omicron strain of SARS-CoV-2 became prevalent, the number of
47	severe cases complicated by pneumonia decreased, and the vaccination had spread socially.
48	In May 2023, the legal classification was changed, and it was decided that a COVID-19
49	infection would be treated on the same level as an influenza virus infection [3].
50	This study was conducted with the aim of determining how the radiology department
51	of a city hospital in Japan could use existing image analysis software to contribute to clinical
52	practice at a time when the pre-Delta strain COVID-19 virus was predominant.
53	During the study period, COVID-19 infection had a high complication rate with
54	pneumonia, especially in older adults [4, 5], with a high rate of aggravation and mortality,
55	and it became necessary to distribute limited medical resources. The discrimination between
56	mild and severe cases at the emergency department was an important and burdensome task.
57	Typically, the severity was determined by symptoms, age, complications, blood tests, and
58	computed tomography (CT) findings. The CT findings were generally evaluated visually, and
59	the CT scores based on visual evaluation were not accurate or objective and took time and
60	effort on the part of the evaluator. There are many reports on the CT severity assessment of
61	COVID-19-associated pneumonia using an imaging software. The measurement methods and
62	evaluation conditions differ for each individual tool, and few of them have been widely
63	adopted in clinical settings.

64 The Ziostation2 image analysis software (Z2; Ziosoft, Tokyo, Japan) had been 65 introduced in approximately 300 facilities in Japan, which was designed to quantify

pulmonary emphysema in patients with chronic obstructive pulmonary disease. When a region above a certain concentration is recognized as a pneumonia region, the pneumonia volume ratio (PVR) can be measured by changing the threshold setting of the CT value (Fig 1a, b). To the best of our knowledge, there are no reports of COVID-19 pneumonia assessment by Z2.

71

Fig 1. Images displayed on the console of the Z2. Z2 monitor screen. The PVR above a
certain concentration is displayed in the upper right corner (red square). LAV, low attenuation
volume; LL, left lower lobe; LU, left upper lobe; nLAV, not LAV (lung volume other than
LAV); PVR, pneumonia volume ratio; RL, right lower lobe; RM, right middle lobe; RU, right
upper lobe; Z2, Ziostation2.

77

Since Z2 has not been set to evaluate pneumonia, it is necessary to determine the threshold in Hounsfield units (HUs) for it. Therefore, it was decided to set the threshold at the concentration that most closely matched the visual evaluation.

In this study, we determined the appropriate conditions for the evaluation of COVID-19 pneumonia by Z2 through comparison with visual evaluation results and examined the usefulness of the clinical severity assessment of Z2.

84

85 Materials and Methods

86 Study population

This study adhered to the principles of the Declaration of Helsinki and was approved by the Ethical Review Committee of Fujisawa City Hospital (approval number: F2021022). The

89	study was conducted retrospectively using imaging data and electronic medical records. An
90	informed consent was provided by all patients in an opt-out manner on the website.
91	We evaluated patients diagnosed with COVID-19 using a polymerase chain reaction
92	test who required a chest CT scan at our hospital and inpatient hospital care between January
93	2020 and January 2021. The patients who had an initial CT scan at another hospital or those
94	who were initially treated at another hospital and subsequently transferred to our hospital, and
95	cases without pneumonia findings on chest CT were excluded. Ten samples were randomly
96	selected from patients under 65 years of age and with an uncomplicated condition.
97	The clinical severity of COVID-19 was classified as mild ($SpO_2 > 93\%$) or severe
98	(SpO ₂ \leq 93%, intubation, and intensive care unit management) based on the symptoms at the
99	time of hospitalization, according to the guidelines of the Ministry of Health and Welfare [4].
100	The clinical severity, symptoms, comorbidities, blood test values, and clinical course were
101	retrieved from the electronic medical records.
102	

103 CT protocol

The chest CT scans were obtained using 64-multidetector CT scanners (SOMATOM Definition AS 64; Siemens Healthineers, Erlangen, Germany). The CT parameters used at our hospital were as follows: 120 kVp, 160-316 mA current intelligent control (auto mA), and 5 mm slice thickness reconstruction. All CT examinations were performed without the use of intravenous contrast agents. The EV Report picture archiving and communication system (PACS) (PSP Corporation, Tokyo, Japan) was used to evaluate the CT findings.

110

111 **CT image analysis**

112 Two radiologists evaluated the CT findings of pneumonia in all patients (Y.N. and M.S.) in 113 consultation for the presence or absence of ground-glass opacity (GGO) (-/+), crazy-paving 114 finding (-/+), consolidation (none/mild/moderate/severe), and emphysema (-/+). 115 For the 10 selected participants, visual evaluation of the PVR was performed independently 116 by two radiologists (Y.F. and M.S.) using the free-form curve drawing tool of the PACS by 117 adding up the area of the lungs and the pneumonia area freehand at 1.5-cm intervals in the 118 coronal chest CT images (Fig 2). In the same participants, the two radiologists independently 119 scored the percentage of pneumonia area in each lobe using visual measurements (0:0%, 1: 120 25%, 2: 25–50%, 3: 50–75%, and 4: 75–100%). 121 122 Fig 2. Visual measurement of PVR. Two radiologists independently selected the entire lung 123 field and pneumonia area every 1.5 cm on the coronal view using a drawing tool on the PACS 124 (PSP Corporation, Tokyo, Japan), and added these up to measure the PVR. The blue line

indicates the entire lung field (mm^2) , and the yellow line indicates the pneumonia area (mm^2) .

126 The minimum and maximum in the figure represent CT values in the region. PACS, report

127 picture archiving and communication system; PVR, pneumonia volume ratio; min, minimum;

128 max, maximum; SD, standard deviation.

129

130 Z2 provided the quantification of the emphysema, healthy lung parenchyma, GGO, 131 and consolidation based on a HU. Z2 can divide segments and calculate total volumes for 132 both the right and left lungs. In the measurement of PVR and CT scores in the 10 selected 133 participants using Z2, the lung fields above a particular concentration were set as pneumonia 134 areas and measured at ≥ -500 HU, ≥ -550 HU, ≥ -600 HU, ≥ -650 HU, and ≥ -700 HU. Z2 135 may not recognize the subpleural consolidation area as a lung field, and the total lung volume

may be underestimated (Fig 3); therefore, radiologist A (M.S.) made the appropriate
corrections manually.

138

Fig 3. Dorsal subpleural consolidations are not recognized as part of the lung and
 require manual correction. The white arrows indicate the areas that needed to be manually
 corrected.

142

143 **Statistical analysis**

144 The presence of significant differences in participant background (age, sex, number of days 145from disease onset to CT evaluation, and laboratory test results) between the mild and severe 146 groups was evaluated using the t-test and chi-square test. The accuracy between the gross 147 measurements of PVR and CT scores by two independent radiologists and the measurements 148 by Z2 were evaluated using the Spearman's rank correlation coefficient. The influence of 149 possible confounding factors of participant background (age, sex, number of days from 150 disease onset to CT evaluation, and presence of comorbidities) on the severity classification 151 of PVR by Z2 was evaluated using the bivariable logistic regression. The usefulness of PVR 152 and CT scores by Z2 under the determined conditions, primary laboratory tests, and CT 153findings in the clinical severity assessment was determined by the receiver operating 154 characteristic (ROC) curves, Youden index, sensitivity, specificity, and p-values. All 155statistical analyses were performed using the SPSS software (version 27; IBM, Armonk, NY, 156 USA).

157

Results

159	The number of patients diagnosed with COVID-19 using a polymerase chain reaction test
160	who required a chest CT scan at our hospital and inpatient hospital care between January
161	2020 and January 2021 were 91. Of these, three patients who received initial treatment at
162	another hospital and 28 patients who had no findings of pneumonia on chest CT were
163	excluded. Two cases were excluded from the study because the thin slice data necessary for
164	Z2 measurement were not saved, and five cases could not be measured by Z2 for unknown
165	reasons.
166	Fig 4. The flow chart shows the process of determining the number of study cases to 53.
167	
168	In total, 53 participants (41 men and 12 women, with a median age of 61.3 years; 38 in the
169	mild group and 15 in the severe group) were included. Table 1 shows the participants'
170	demographics (age, sex, and presence of comorbidities), laboratory findings, and CT findings.
171	Fifty-two participants presented with COVID-19 symptoms; however, there was no
172	significant difference in the severity of the symptoms between the mild and severe disease
173	groups. Significant differences in the number of days from disease onset to CT evaluation
174	and the presence of comorbidities were found between the two groups. In addition, laboratory
175	results revealed that C-reactive protein (CRP) and lactate dehydrogenase (LDH) levels
176	differed significantly between the two groups. The CT findings showed a significant
177	difference in consolidation between the two groups.
178	

179 Table 1. Patient background, blood test, and CT findings

FactorTotal (n = 53)Mild (n = 38)Severe (n = 15)p-value

Age (years; median)	61.28 (66)	58.95 (59.5)	67.2 (67.00)	0.148
Sex (male; %)	41 (77.4)	27 (71.1)	14 (93.3)	0.081
Date from onset to	$\epsilon_{0}(1,14)$	5 2 (1 12)	70(4,14)	0.016
CT (range)	0.0 (1-14)	3.2 (1-12)	7.9 (4–14)	0.010
Comorbidities ^a (%)	18 (34.0)	9 (23.7)	9 (60.0)	0.012
DM (%)	6 (11.3)	3 (7.9)	3 (20.0)	0.21
COPD (%)	5 (9.4)	2 (5.3)	3 (20.0)	0.098
CRF (%)	4 (7.5)	3 (7.9)	1 (6.7)	0.879
Obesity (%)	6 (11.3)	1 (2.6)	5 (33.3)	0.0015
Malignancy (%)	3 (5.7)	1 (2.6)	2 (13.3)	0.129
Symptoms (any)	52 (98.1)	37 (97.4)	15 (100)	0.526
Fever (%)	46 (86.8)	33 (86.8)	13 (86.7)	0.986
Cough (%)	22 (41.5)	14 (36.8)	8 (53.3)	0.272
Taste disorder (%)	7 (16.2)	7 (18.4)	0 (0)	0.074
Vomiting or diarrhea	0(17.0)	8 (21.1)	1 (67)	0.21
(%)	9 (17.0)	8 (21.1)	1 (0.7)	0.21
Blood tests (range)				
WBC (×10 ⁹ /L)	6.6 (2.5–21.7)	6.3 (2.5–21.7)	7.2 (3.5–12.3)	0.33
Lymphocytes (%)	17.8 (4–46.5)	11.6 (4–46.5)	7.5 (5.2–26.4)	0.085
CRP (mg/dL)	8.1 (0.02–29.6)	2.9 (0.02–29.6)	13.1 (2.4–26.0)	0.005
	364.8	302.8	501.9 (022, 1127)	< 0.001
LDH(0/L)	(144–1136)	(144–834)	521.8 (255–1150)	< 0.001
AST (U/L)	50.2 (14–160)	46.3 (14–160)	60.1 (25–131)	0.213
ALT (U/L)	43.5 (7–200)	39.97 (7–200)	52.27 (15–163)	0.308
Creatinine (mg/dL)	1.39 (0.31–14.9)	1.15	0.998 (0.52–2.34)	0.455

		(0.31–14.9)		
eGFR (mL/min)	65.8 (30–144)	65.4 (1.6–144)	66.93 (22–100)	0.853
CT findings				
GGO (+) (%)	53 (100)	38 (100)	15 (100)	1
Crazy paving (+) (%)	8 (15.1)	5 (13.2)	3 (20.0)	0.53
Consolidation (-)	22(42.4)	21(55,2)	2(12.2)	
(%)	23 (43.4)	21 (33.3)	2 (13.3)	
(+)(%)	17 (32.1)	12 (31.6)	5 (33.3)	
(++)(%)	8 (15.1)	5 (13.2)	3 (20.0)	
(+++)(%)	5 (9.4)	0 (0)	5 (33.3)	0.0006
Z2 (≥ –600 HU)				
PVR mean (median,	12.44	7.59	24.71	< 0.001
range)	(1.63–63.26)	(1.63–40.11)	(3.32–63.26)	< 0.001
CT score mean	6 62 (9 5 15)	5 97 (5 5 11)	9 52 (7 5 15)	< 0.001
(median, range)	0.02 (0, 3–13)	3.07 (3, 3-11)	0.33 (7, 3-13)	< 0.001

180

CT: computed tomography, DM: diabetes mellitus, COPD: chronic obstructive pulmonary disease, CRF: chronic renal failure, WBC: white blood cell, CRP: C-reactive protein, LDH: lactate dehydrogenase, AST: aspartate aminotransferase, ALT: alanine aminotransferase, eGFR: estimated glomerular filtration rate, GGO: ground-glass opacity, Z2: Ziostation2, PVR: pneumonia volume ratio

^aComorbidities were defined as presence of any of the following: DM, COPD, severe cardiovascular disease, severe CRF, obesity, malignancy under treatment, immunosuppression, and liver cirrhosis.

189

190 Table 2 shows the results of the Spearman's correlation between Z2 (under each 191 condition; PVR: \geq -500 HU, \geq -550 HU, \geq -600 HU, \geq -650 HU, and \geq -700 HU, CT score: 192 \geq -500 HU, and \geq -600 HU) and the two radiologists for PVR and the CT scores in the 10 193 participants without comorbidities, respectively. While the accuracy between the two 194 radiologists and Z2 for PVR was equally high at ≥ -500 HU to ≥ -600 HU, the accuracy for 195 CT scores was higher at \geq -600 HU than at \geq -500 HU. Based on these results, the Z2 196 measurement condition for COVID-19 pneumonia that achieved the best accuracy with the 197 gross measurement was determined to be ≥ -600 HU.

198

Table 2. Results of the Spearman's test of PVR and CT score by two radiologists and
 Ziostation2 of five/two conditions in the 10 selected patients

		Reader	≥-500	≥-550	≥-600	≥-650	≥-700
		В	HU	HU	HU	HU	HU
Reader A	PVR	0.879	0.976	0.976	0.976	0.964	0.818
	CT score	0.976	0.639		0.651		
Reader B	PVR		0.842	0.842	0.842	0.83	0.661
	CT		0.594		0 606		
	score		0.584		0.696		

201

202 PVR: pneumonia volume ratio, CT: computed tomography,

203 Reader A: M.S., Reader B: Y.F.

205	Figs 4 and 5 show the ROC curves and boxplots corresponding to the classification
206	of disease severity by PVR and CT score using Z2 (\geq -600HU), CRP, and LDH. The areas
207	under the curve (AUCs) were 0.881, 0.77, 0.788, and 0.842, respectively.
208	
209	Fig 5. ROC curves for PVR, CT score, CRP, and LDH. ROC curve for a . PVR using Z2 (\geq
210	–600 HU) and b . CT scores using Z2 (\geq –600 HU), c . CRP, and d . LDH. ROC, receiver
211	operating characteristic; PVR, pneumonia volume ratio; Z2, Ziostation2; CT, computed
212	tomography; CRP, C-reactive protein; LDH, lactate dehydrogenase; SD, standard deviation
213	
214	Fig 6. Boxplots of PVR, CT score, and CRP. Boxplots for a. PVR using Z2 (\geq -600 HU), b.
215	CT scores using Z2 (\geq -600 HU), c . CRP, and d . LDH. 1: mild group, 2: severe group. Error
216	bars indicate outliers. PVR, pneumonia volume ratio; CT, computed tomography; Z2,
217	Ziostation2; CRP, C-reactive protein; LDH, lactate dehydrogenase
218	
219	The Youden index values for PVR and CT scores at \geq -600 HU by Z2, CRP, and
220	LDH were 18.69, 7.5, 5.26, and 306.5, respectively. The sensitivities for PVR and CT scores
221	at \geq –600 HU by Z2 were 66.7% and 40%, respectively. The specificities for PVR and CT
222	scores at \geq –600 HU by Z2 were 94.7% and 74%, respectively. The p-value for PVR at \geq
223	–600 HU by Z2 was $p < 0.001,$ and that for CT scores at \geq – 600 HU by Z2 was $p = 0.002$
224	(Table 3). The bivariable logistic regression of PVR (≥–600 HU) according to age, sex, date
225	from onset to CT, and comorbidities showed no significant effects, except for comorbidities
226	(Table 4). The sensitivity and specificity were 66.7% and 89.5% when the PVR threshold
227	was 18, and 60% and 97.4% when the PVR threshold was 20, respectively.

228

Table 3. Cut-off values for pneumonia volume ratio and blood test to differentiate mild

and severe groups

	Youden	AUC	Sensitivity	Specificity	Lower 95%	Upper 95% CI	p-value
					CI		
PVR (≥ –600	18 69	0.881	66.7	94 7	0 781	0.981	< 0.001
HU)	18.09	0.001	00.7	21.7	01701	0.901	
CT score (≥	75	0.77	40	74	0.620	0.011	0.002
-600 HU)	7.5	0.77	40	/4	0.029	0.911	0.002
CRP	5.26	0.788	86.7	68.4	0.664	0.912	< 0.001
LDH	306.5	0.842	86.7	68.4	0.729	0.956	< 0.001

231

AUC: area under the curve, PVR: pneumonia volume ratio, CT: computed tomography, CRP:

233 C-reactive protein, LDH: lactate dehydrogenase, CI: confidence interval

234

Table 4. Bivariable logistic regression of PVR (≥–600 HU) according to age, sex, number

236 of days from onset to CT, and comorbidities

Predictor	OR (95% CI)	p-value
PVR (≥ –600 HU)	1.131 (1.048–1.221)	0.002
Age	1.031 (0.979–1.086)	0.246
PVR (≥ -600 HU)	1.124 (1.049–1.206)	0.001
Sex	0.000 (0.146–1530.796)	0.252

PVR (≥ –600 HU)	1.126 (1.044–1.214)	0.002
Number of days from onset to CT	1.058 (0.837–1.339)	0.637
PVR (≥ –600 HU)	1.137 (1.045–1.237)	0.003
Comorbidities (any)	9.795 (1.432–67.002)	0.02

238

PVR, pneumonia volume ratio; OR, odds ratio; CT, computed tomography; CI, confidence
interval

241

The evaluation of PVR and CT scores in patients affected by COVID-19-associated pneumonia by Z2 was highly consistent with the visual-evaluation results under the condition of \geq -600 HU. The AUC and Youden index of the ROC curve by Z2 (\geq -600HU) were 0.881 and 18.69 for PVR, and 0.77 and 7.5 for the CT score, respectively, indicating that they are useful for clinical severity classification.

247

248 **Discussion**

The chest CT plays a major role in COVID-19 treatment, including severity judgment and prognostic prediction. In clinical practice and in previous studies, the spatial progression of pneumonia on CT has been evaluated with naked eye, and the accuracy and homogeneity have not been ensured.

In this study, we examined the usefulness of determining the severity of COVID-19-associated pneumonia using Z2, an image analysis software widely available in

Japan. This methodology can be easily deployed at facilities that have Z2 and thus has high clinical utility.

Several reports evaluated the percentage of lesion area of COVID-19-associated pneumonia 257 258 in each lobe of the lung visually and scored them to determine the disease severity [6-10]. 259 Yang et al. [6] visually classified the percentage of lesion area in each segment as 0%, <50%, 260 and >50%. Li et al. [7] reported that the percentage of lesion area in each lobe was visually 261 classified as 0%, 0–25%, 25–50%, 50–75%, and 75–100%, and scored on a scale of 0–20. 262 The authors found that the optimal threshold for the severe group was 7.5. Francone et al. [9] 263 used a similar classification, with a mortality risk cut-off of 18. Li et al. [8] also reported 264 scores of 0:0%, 1:<5%, 2:5–25%, 3:25–50%, 4:50–75%, and 5:>75% or higher, with a 265 cut-off score of 7, a sensitivity of 80%, and a specificity of 82.8% for the severely ill group. 266 The cut-off value for clinical severity classification by CT score varies depending on the 267 method and on how the severity is classified.

268 The CT scores based on visual evaluations that do not require special software or techniques 269 are widely used in clinical settings. This type of evaluation is subjective; however, it has been 270 reported that the inter-evaluator difference is small, and the results of this study are in 271 agreement. However, the score measurement for each lobe in 25-50% increments is 272 troublesome and imposes a burden on the emergency unit staff. Inoue et al. [11] reported that 273 three visual CT score evaluations required 25.7-41.7 s, 27.7-39.5 s, and 48.9-80.0 s, 274 respectively. Novel methods for the quantitative and automated measurement of the spatial 275 progression of COVID-19-associated pneumonia have been reported since the early days of 276 the pandemic [12-21].

Using the commercially available image analysis software, Timaran-Montenegro et al. [12] automatically classified –700 to –1000 HU as normal lung, and –500 to 20 HU as pneumonia regions, and compared the survival vs non-survival groups. The percentage of normal lung

was a significant independent factor according to a multinominal logistic analysis. Colombi et al. [13] defined the region of -950 to -700 HU as well as aerated lungs and reported that the measurement by commercially available software and visual measurement were very similar and useful for severity evaluation. In the 10 cases selected in our study, the correlation between the automated measurement by Z2 under the condition of ≥ -600 HU and the macroscopic measurement was high: very high for PVR (correlation coefficient 0.842-0.976) and moderate for the CT score (correlation coefficient 0.651-0.696).

As there were no previous reports of using Z2 as a tool to evaluate diseases such as pneumonia with increased lung concentration, the concentration range for pneumonia was determined to be \geq -600 HU in this study, based on the high degree of consistency in terms of visual PVR and CT score.

291 The range of normal lung, GGO, and consolidation reported in each study using software

varied as follows: between: -1000 to -600 HU for normal lung, -750 to -100 HU for GGO,

and -399 to -69 HU for consolidation [10-15]. Many previous studies set the lower limit of

the GGO range at -800 to -700 HU; however, -600 HU was selected as the lower limit in

this study due to the high degree of agreement with the visual findings. This was probably

because it is difficult to recognize a faint increase in concentration based on visual evaluation

297 compared to the software-assisted evaluation. It is an advantage of the software-assisted

evaluation that it can detect faint concentrations; however, considering that the CT evaluation

of COVID-19-associated pneumonia is generally based on visual evaluation, the detection of

300 faint concentrations that are not measurable by visual evaluation leads to clinical

301 discrepancies.

Grassi et al. [14] reported that the percentage of normal lung, emphysema, and consolidation measured by three different software tools were inconsistent. Granata et al. [15] compared the results obtained from two different software tools and reported that the correlation between

them was not high enough. The algorithms in which each software is based are different, and therefore comparisons cannot be made under uniform conditions. Z2 is a software tool owned by more than 300 facilities in Japan. Therefore, an assessment method based on the use of Z2 may be immediately available at these facilities and have a high clinical significance. In addition, the introduction of new technologies is time-consuming and expensive.

Okuma et al. [17] reported that the CT score and the percentage of opacity (PVR in this study) obtained using commercially available AI-based software showed a similar AUC; however, in this study the AUC corresponding to PVR and the CT score estimated by Z2 under \geq -600 HU was higher in the case of PVR. Theoretically, the CT scoring method can differ by up to 24% in one lobe at the same point, making it less accurate than PVR. When automated measurement of the same standard becomes widespread, the evaluation by PVR is likely to replace CT scores.

Recently, there have been many reports on the diagnosis and severity assessment of COVID-19-associated pneumonia using AI [16-20]. In a study on COVID-19-associated pneumonia using an AI-based software developed by Ziosoft, the company that developed Z2, Aoki et al. [20] measured the CT lesion extent separately for normal lung, GGO, reticulation, and consolidation. In this study, the pneumonia area was evaluated by combining GGO and consolidation; however, more accurate qualitative and quantitative evaluation will be possible if AI-based software is adopted for this purpose in the future.

In this study, Z2 sometimes misidentified subpleural consolidation as extrapulmonary, requiring manual correction. Inoue et al. [11] reported the measurement errors with the use of U-NET due to the inclusion of atelectasis, fibrosis, and air trapping in the density mask. When a software tool is used, the measurement is carried out automatically; however, the error checking may still need to be performed by human staff.

In this study, we showed the optimal conditions for measuring the PVR and CT score in cases of COVID-19-associated pneumonia using Z2, a widely used image analysis software in Japan, and provided a guideline for clinical severity evaluation based on it. Therefore, defining a Z2-based assessment method has a high clinical significance, and replacing visual evaluation with existing image analysis software represents a way to quickly reduce the burden on clinicians at each facility.

Binomial logistic regression analysis showed no significant effects of age, sex, or time from onset to CT on PVR.

In terms of CT findings, consolidation was significantly higher among the severe group, in agreement with previous reports [9, 19-21]. Several laboratory tests have been reported to be indicators of COVID-19 infection. In our study, both CRP and LDH were significant items, again in agreement with previous reports [22, 23].

The major limitation of this study was the small number of participants at a single facility. The other limitations were that the manual correction of the subpleural consolidation in the Z2 measurement was performed by a single radiologist and the significance of inter-operator differences was not evaluated. Moreover, PVR assumed the area of \geq -600 HU to be a surrogate value for COVID-19 pneumonia, but no histological confirmation was available. The PVR measurements were uniformly performed regardless of the background lesions affecting emphysema, fibrosis, or atelectasis.

In conclusion, we determined the optimal conditions that best approximates visual evaluation for assessing COVID-19-associated pneumonia using Z2, one of the most popular image analysis software tools in Japan and demonstrated that the AUC of PVR was higher than that of CT score in the assessment of clinical severity. The introduction of new technologies is time-consuming and expensive; this method has high clinical utility and can be adopted immediately in any facility where Z2 is available for use.

354

355 Acknowledgments

356	We would like to thank Dr. Noriko Hida and Dr. Eisuke Inoue for their guidance on the
357	statistical analysis, and Ms. Hokazono and Editage (www.editage.com) for English language
358	editing. We also appreciate the support from our proofreaders and editors. In addition, we are
359	grateful to the clinicians at Fujisawa City Hospital for their insightful advice.
360	
361	
362	
363	
364	
365	
366	
367	
368	
369	
370	
371	
372	

3	7	3
J		υ

374

375

376 **References**

377	1. Hui DS, Az	zhar EI, Madar	i TA, Ntou	mi F, Kock R	, Dar O, et	al. The continuing
-----	---------------	----------------	------------	--------------	-------------	--------------------

- 2019-nCoV epidemic threat of novel coronaviruses to global health—The latest 2019
- novel coronavirus outbreak in Wuhan, China. Int J Infect Dis. 2020;91: 264-266. doi:
- 380 10.1016/j.ijid.2020.01.009.
- 2. Furuse Y, Ko YK, Saito M, Shobugawa Y, Jindai K, Saito T, et al. Epidemiology of
- 382 COVID-19 outbreak in Japan, from January-March 2020. Jpn J Infect Dis. 2020;73:
- 383 **391-393.** doi: <u>10.7883/yoken.JJID.2020.271</u>.
- 384 3. Japanese Ministry of Health, Labour and Welfare. Regarding the response after the
- transition to Category 5 infectious diseases of the new coronavirus infectious disease;

386 2023. Available from: https://www.mhlw.go.jp/stf/corona5rui.html.

- 4. Japanese Ministry of Health, Labour and Welfare. COVID-19 in Japan. Available from:
 https://covid19.mhlw.go.jp/extensions/public/index. html (2022).
- 5. Matsunaga N, Hayakawa K, Terada M, Ohtsu H, Asai Y, Tsuzuki S, et al. Clinical
- epidemiology of hospitalized patients with coronavirus disease 2019 (COVID-19) in
- Japan: report of the COVID-19 Registry Japan. Clin Infect Dis. 2021;73: e3677-e3689.
- 392 doi: <u>10.1093/cid/ciaa1470</u>.
- 6. Yang R, Li X, Liu H, Zhen Y, Zhang X, Xiong Q, et al. Chest CT severity score: an
- imaging tool for assessing severe COVID-19. Radiol Cardiothorac Imaging. 2020;2:
- e200047. doi: <u>10.1148/ryct.2020200047</u>.

- ³⁹⁶ 7. Li K, Fang Y, Li W, Pan C, Qin P, Zhong Y, et al. CT image visual quantitative evaluation
- and clinical classification of coronavirus disease (COVID-19). Eur Radiol. 2020;30:
- 398 4407-4416. doi: <u>10.1007/s00330-020-06817-6</u>.
- 8. Li K, Wu J, Wu F, Guo D, Chen L, Fang Z, et al. The clinical and chest CT features
- 400 associated with severe and critical COVID-19 pneumonia. Invest Radiol. 2020;55:
- 401 327-331. doi: <u>10.1097/RLI.00000000000672</u>.
- 402 9. Francone M, Iafrate F, Masci GM, Coco S, Cilia F, Manganaro L, et al. Chest CT score in
- 403 COVID-19 patients: correlation with disease severity and short-term
- 404 prognosis. Eur Radiol. 2020;30: 6808-6817. doi: <u>10.1007/s00330-020-07033-y</u>.
- 10. Liang T, Liu Z, Wu CC, Jin C, Zhao H, Wang Y, et al. Evolution of CT findings in
- 406 patients with mild COVID-19 pneumonia. Eur Radiol. 2020;30: 4865-4873. doi:
- 407 <u>10.1007/s00330-020-06823-8</u>.
- 11. Inoue A, Takahashi H, Ibe T, Ishii H, Kurata Y, Ishizuka Y, et al. Comparison of
- 409 semiquantitative chest CT scoring systems to estimate severity in coronavirus disease
- 410 2019 (COVID-19) pneumonia. Eur Radiol. 2022;32: 3513-3524. doi:
- 411 <u>10.1007/s00330-021-08435-2</u>.
- 412 12. Timaran-Montenegro DE, Torres-Ramírez CA, Morales-Jaramillo LM, Mateo-Camacho
- 413 YS, Tapia-Rangel EA, Fuentes-Badillo KD, et al. Computed tomography-based lung
- 414 residual volume and mortality of patients with coronavirus disease-19 (COVID-19). J
- 415 Thorac Imaging. 2021;36: 65-72. doi: <u>10.1097/RTI.00000000000572</u>.
- 416 13. Colombi D, Bodini FC, Petrini M, Maffi G, Morelli N, Milanese G, et al. Well-aerated
- 417 lung on admitting chest CT to predict adverse outcome in COVID-19
- 418 pneumonia. Radiology. 2020;296: E86-E96. doi: <u>10.1148/radiol.2020201433</u>.
- 419 14. Grassi R, Cappabianca S, Urraro F, Feragalli B, Montanelli A, Patelli G, et al. Chest CT
- 420 computerized aided quantification of pneumonia lesions in COVID-19 infection: A

421	comparison	among three	commercial	software.	Int J	Environ	Res P	ublic	Health.
101	companioon	unions unice	commercial	bon marc.	IIIC 0 1	Lin in on .		aone	ricultil.

- 422 **2020;17: 6914. doi:** <u>10.3390/ijerph17186914</u>.
- 423 15. Granata V, Ianniello S, Fusco R, Urraro F, Pupo D, Magliocchetti S, et al. Quantitative
- 424 analysis of residual COVID-19 lung CT features: consistency among two commercial
- 425 software. J Pers Med. 2021;11: 1103. doi: <u>10.3390/jpm11111103</u>.
- 426 16. Kauczor HU, Heitmann K, Heussel CP, Marwede D, Uthmann T, Thelen M. Automatic
- 427 detection and quantification of ground-glass opacities on high-resolution CT using
- 428 multiple neural networks: comparison with a density mask. AJR Am J Roentgenol.
- 429 2000;175: 1329-1334. doi: <u>10.2214/ajr.175.5.1751329</u>.
- 430 17. Okuma T, Hamamoto S, Maebayashi T, Taniguchi A, Hirakawa K, Matsushita S, et
- 431 al. Quantitative evaluation of COVID-19 pneumonia severity by CT pneumonia
- 432 analysis algorithm using deep learning technology and blood test results. Jpn J Radiol.
- 433 2021;39: 956-965. doi: <u>10.1007/s11604-021-01134-4</u>.
- 434 18. Durhan G, Ardalı Düzgün S, Başaran Demirkazık F, Irmak İ, İdilman İ, Gülsün Akpınar
- 435 M, et al. Visual and software-based quantitative chest CT assessment of COVID-19:
- 436 correlation with clinical findings. Diagn Interv Radiol. 2020;26: 557-564. doi:
- 437 <u>10.5152/dir.2020.20407</u>.
- 438 19. Huang L, Han R, Ai T, Yu P, Kang H, Tao Q, et al. Serial quantitative chest CT
- 439 assessment of COVID-19: A deep learning approach. Radiol Cardiothorac Imaging.
- 440 2020;2: e200075. doi: <u>10.1148/ryct.2020200075</u>.
- 441 20. Aoki R, Iwasawa T, Hagiwara E, Komatsu S, Utsunomiya D, Ogura T. Pulmonary
- 442 vascular enlargement and lesion extent on computed tomography are correlated with
- 443 COVID-19 disease severity. Jpn J Radiol. 2021;39: 451-458. doi:
- 444 <u>10.1007/s11604-020-01085-2</u>.

- 445 21. Li Z, Zhong Z, Li Y, Zhang T, Gao L, Jin D, et al. From community-acquired pneumonia
- 446 to COVID-19: a deep learning-based method for quantitative analysis of COVID-19 on
- thick-section CT scans. Eur Radiol. 2020;30: 6828-6837. doi:
- 448 <u>10.1007/s00330-020-07042-x</u>.
- 449 22. Shi H, Han X, Jiang N, Cao Y, Alwalid O, Gu J, et al. Radiological findings from 81
- 450 patients with COVID-19 pneumonia in Wuhan, China: a descriptive
- 451 study. Lancet Infect Dis. 2020;20: 425-434. doi: <u>10.1016/S1473-3099(20)30086-4</u>.
- 452 23. Kurashima K, Kagiyama N, Ishiguro T, Kasuga K, Morimoto Y, Ozawa R, et
- 453 al. Predictors of severe COVID-19 pneumonia. Kansenshogaku Zasshi. 2020;94:
- 454 483-489. doi: <u>10.11150/kansenshogakuzasshi.94.483</u>.



/	LAV %	nLAV	nLAV%
2ee	9192%	33590cc	8.095
300	9037%	187,2300	9.635
600	89.68%	116/19cc	10.225
7cc	9133%	7C.74cc	8,675
9cc	9020%	140.66oc	6.725
6cc	92.75%	38.13oc	7255
7cc	95.57%	292900	4,435
7ee	92.09%	812500	7515





Fig 4.







CRP



LDH



