CT-severity Score in COVID-19 patients, Assessment of Performance in Triage and Outcome Prediction: a Comparative Study of Different Methods

Alireza Almasi Nokiani[1]

[1]Affiliation not available

December 24, 2021

Abstract

BACKGROUND: Lung involvement in COVID-19 can be quantified by chest CT scan with some triage and prognostication value. At least 7 CT-severity score (CTSS) systems have been proposed. **PURPOSE:** We evaluated triage and prognostication performance of seven different CTSSs for COVID-19. **MATERIALS AND METHODS:** COVID-19, PCR positive patients admitted from February 20th 2020 to July 22nd were included into a retrospective study. Demographic data and clinical data indicating disease severity at presentation and in peak disease severity were recorded. CT images were reviewed and scored according to seven different scoring systems (CTSS1-CTSS7) by two radiologists. Interrater reliability was determined for each CTSS. Then clinical severity of the disease at presentation (for triage) and peak disease severity (for outcome) were compared with CTSSs separately. ROC curves for performance of each CTSS in diagnosing severe/critical disease on admission, severe/critical disease at peak disease severity and critical disease at peak severity were plotted. Areas under the curve (AUCs), best thresholds and corresponding sensitivities and specificities were calculated. **RESULTS:** 96 patients were included with mean age of 63.6 ± 17.4 years (range: 21-88, median: 67). 57 (59,4%) were men and 39 (40.6%) were women. All CTSSs showed good interrater reliability as calculated intraclass correlation coefficients (ICCs) were 0.764-0.837 for all of the CTSSs. Only three CTSSs showed acceptable AUCs (AUC =0.7) for triage of severe/critical patients. All CTSSs showed acceptable AUCs for prognostication (AUCs=0.76-0.79). Calculated AUCs were not significantly different for triage and for prediction of severe/critical disease but some difference was shown for prediction of critical disease. **CONCLUSION:** Men are probably affected more frequently than women by COVID19. CTSS performance in triage was much lower than earlier reports and only three CTSSs showed acceptable AUCs. CTSS performed better for prognostic purposes than for triage as all 7 CTSSs showed acceptable AUCs in both types of prognostic ROC curves. Our results are compatible with those of recent studies. There is not much difference among performance of different CTSSs.

Authors

1. Alireza Almasi Nokiani MD Interventional Radiologist[a,b] Email: almasinokiani.ar@iums.ac.ir
2. Razieh Shahnazari MD Radiologist [a,b] Email: rsh.medicen1999@gmail.com
3. Mohammad Amin Abbasi MD Internist [b] Email: amin.abbasi1314@gmail.com
4. Farshad Divsalar MD Infectious Disease Specialist [b]Email: divsalar.f@iums.ac.ir
5. Marzieh Bayazidi MD Radiology Resident [a] Email: mar.baya.dr@gmail.com
6. Azadeh Sadatnaseri MD Cardiologist [c] Email: asn350@yahoo.com

Institutions

1. Department of Radiology, Firoozabadi Hospital, Iran University of Medical Sciences, Tehran, Iran
2. Firoozabadi Clinical Research Development Unit (FCRDU), Iran University of Medical Sciences, Tehran, Iran

3. Department of Cardiology, Sina Hospital, Tehran University of Medical Sciences, Tehran, Iran

Corresponding Author:

Azadeh Sadatnaseri

Original research

List of abbreviations:

AUC : Area Under the Curve

CI: Confidence Interval

CTSS : Computed Tomography Severity Score

HIS : Hospital Information System

ICC : Inraclass Correlation Coefficient

PACS : Picture Archiving and Communication System

ROC curve : Receiver Operator Characteristic curve

RT-PCR : Reverse-Transcriptase Polymerase-Chain-Reaction

Introduction

Because of primary involvement of respiratory system, chest CT is recommended in suspected COVID-19 cases [1]. Lung involvement in COVID-19 can be quantified by chest CT with triage and prognostication value [1-10]. We aimed to determine the value of CTSS in making decisions about the intensity of the treatment of respiratory failure (triage) and predicting the risk of development of severe/critical disease in the course of COVID-19 in correlation with selected clinical parameters (prognostication).

Xie and colleagues used a CT severity score (CTSS) based on dividing the lungs into upper, middle and lower zones, each scored 0-4 according to percentage of involvement (CTSS1) [2]. They stated elsewhere that mean CTSS1 was significantly higher in severe/critical group than in mild/moderate group of patients (12.86 vs 5.34) [3]. Zhou and co-workers used a CTSS with the same zonal concept, further dividing each zone into anterior and posterior divisions with maximum 48 scores (CTSS2) [4]. There was no performance report. Chung and colleagues scored each of the five lung lobes by percentage of involvement from 0-4. CTSS was the sum of the five lobe scores, with a maximum of 20 (CTSS3) [5]. Li and colleagues implemented CTSS3 and reported an intraclass correlation coefficient (ICC) of 0.976 between two observers and area under the curve (AUC) of 0.918 for receiver operator characteristic (ROC) curve to diagnose severe/critical disease; the CTSS cut-off point of 7.5 had 82.6% sensitivity and 100% specificity [6]. Other researchers used another CTSS. Each of the 5 lung lobes was visually scored from 0 to 5 as: 0, no involvement; 1, <5%; 2, 5-25%; 3, 26-49%; 4, 50%-75% and 5, >75% involvement. Maximum total score was 25 (CTSS4) [7,8]. They reported no ROC curve or cut-off point. We propose another CTSS which is almost the same as CTSS4, but considers lingula as a separate lobe (CTSS5) with a maximum score of 30. Xiong and co-workers assessed each lobe for opacification and lesion size with a maximum sum of 20 (CTSS6) [9]. Yang and colleagues developed another CTSS in which the 18 segments of the lung were divided into 20 regions. The lung opacities in all the 20 lung regions were evaluated on chest CT using a system attributing scores of 0, 1, and 2 according to absence or presence of 50% or more segmental opacification with a maximum of 40 (CTSS7). Interrater reliability for CTSS7 was excellent (ICC=0.936). The area under the ROC curve for diagnosing patients in severe/critical group was 0.892 (95% confidence interval: 0.814-0.944). Optimal CTSS threshold for identifying severe/critical patients was 19.5, with 83.3% sensitivity and 94% specificity. The interrater reliability for CTSS7 was excellent (ICC_{median}=0.925, ICC_{mean}=0.936) [10].

Materials and Methods

Patients

Our institutional review board waived requirement to obtain written informed consent for this retrospective study, which evaluated de-identified data and involved no potential risk for patients. To avert any potential breach of confidentiality, no link between the patients and the researchers was made available.

We enrolled patients with COVID-19 referred to Firoozabadi hospital from February 20th 2020 to July 22nd. The diagnosis was based on positive results of reverse-transcriptase polymerase-chain-reaction (RT-PCR) assay of nasal and pharyngeal swab specimens at any time during hospitalization. Exclusion criteria were significant cardiopulmonary comorbidity, defined as cardiothoracic ratio >60% on CT topogram image [11] and diameter ratios of central branches of pulmonary artery to corresponding bronchi >2 [12,13] or preexisting pulmonary disease involving more than 30% of the lungs, diagnosed subjectively by visual assessment of the same CT images by the radiologist (AA). Patients that did not have any CT examination in our hospital were also excluded.

We retrospectively collected clinical and laboratory data from the hospital information system (HIS), including disease severity at presentation, severity in the most severe disease period, final outcome (death or discharge), place of hospital admission (ward or ICU), state of intubation and any comorbidity.

Severity of the disease was decided by the information derived from patients' records as is presented in table 1 [14]. For less complexity when the exact required data were not available, we regarded those who had undergone tracheal intubation or had died from the disease as critical.

Image acquisition

Chest CT imaging was performed by a 16-detector CT scanner (Emotion; Siemens; Germany). All patients were examined in supine position. CT images were then acquired during a single inspiratory breath-hold. The scanning range was from the apex of lung to costophrenic angle. CT scan parameters: X-ray tube parameters - 110KVp, 45-60 effective mAs; rotation time - 0.6 second; collimation- 16x1.2; pitch - 1.5; section thickness – 5 mm; reconstruction interval – 5 mm with B70 sharp convolution kernel; additional reconstructions at slice thickness, and reconstruction interval of 1.5 mm with B70 and B31 convolution kernels were also made to generate lung and mediastinal windows, respectively. Lung window images were viewed at a width/level of 1200/-600 and mediastinal window images at 350/50 window settings.

Image interpretation

Two radiologists with 17 and 3 years of experience (AAN and RSh respectively) blinded to clinical data reviewed CT images of all the patients independently and scored each patient's images according to each of the 7 scoring systems mentioned in the introduction section (table 2). They viewed images on hospital PACS (Marco PACS Version 2.0.0.0) and resorted to multiplanar reconstruction (MPR) whenever needed. We took into account 11 of 14 imaging features defined in a previous study [15]: ground-glass opacitiy (GGO), consolidation, mixed GGO and consolidation, centrilobular nodules, architectural distortion, tree-in-bud, bronchial wall thickening, reticulation, subpleural bands, traction bronchiectasis and vascular enlargement in the lesion. Other relevant pathological findings such as enlarged heart, other pulmonary parenchymal disease such as cavities and emphysema, pleural effusion and mediastinal lymph nodes were also recorded.

Statistical analysis

All statistical analyses were done using SPSS 26.0 software (IBM, Armonk, NY), excluding comparison of ROC curves and AUCs and selection of cut-off points which were conducted by MedCalc statistical software version 19.9.4.0. $P<0.05$ was considered statistically significant. Statistical analysis was performed by AAN. Quantitative data were expressed as mean ± standard deviation and/or median. Comparison of means was performed by independent-sample t-test for two means and ANOVA test for more than two means [16]. Interrater reliability was evaluated using intraclass correlation coefficient (ICCs) for CTSSs. ICC estimates and their 95% confidence intervals (CI) were calculated based on a two-way random model, single

measurement form and, absolute agreement type ($ICC_{1,1}$ with absolute agreement) [17]. ICCs were classified as follows: poor reliability <0.5; moderate reliability, 0.5–0.74; good reliability, 0.75–0.89; and excellent reliability, 0.9–1.0) [18]. ROC curve analysis was performed on the averages of reported CTSSs by the two raters for each CTSS to calculate AUC for diagnosing severe/critical COVID-19 at the time of hospital admission (for triage). Then AUCs were classified unsatisfactory if AUC<0.7, acceptable if 0.7[?]AUC<0.8, excellent if 0.8[?]AUC<0.9 and outstanding if AUC[?]0.9 [19]. If AUC was acceptable or better, threshold, specificity and sensitivity for the CTSS was calculated. We chose best thresholds according to Youden index method which is choosing the threshold producing the largest Youden Index (sensitivity+ specificity -1) [20]. The AUCs for the ROC curves were compared pairwise by the z test.

The same statistical procedure was applied to the CTSSs for predicting severe/critical disease at peak disease severity and also for predicting critical disease at peak severity (for prognostication).

Results

Among COVID19 patients who referred to our hospital from February 20th 2020 to July 22nd, there were 145 confirmed cases. Of these patients, 110 have had at least one CT scan record in the hospital PACS. After reviewing the first CT images, 14 patients with cardiopulmonary comorbidity were excluded, consisting of 13 patients with significant heart failure and one patient with significant centrilobular emphysema. 96 patients were included in the study. Patient selection process is summarized in figure 1.

In the study group, the mean age was 63.6 +- 17.4 years (range: 21-88 years, median: 67). 57 (59,4%) were men and 39 (40.6%) were women. Disease severity at the time of hospitalization was as follows: 41 (42.7%) moderate, 53 (55.2%) severe and 2 (2.1%) critical. In the most severe period of their disease 22 (22.9%) were moderate, 31 (32.3%) severe and 43 (44.8%) critical. 40 (41.7%) patients died. Demographic and clinical data are summarized in table 3.

All 96 patients underwent initial thoracic CT scan within first 24 hours of admission, on average 4+-3.4 days (range 0-19 days, median 3 days) after the onset of symptoms.

Inter-rater reliabilities between two raters for CTSSs 1-7 calculated as ICCs, as well as related inference, is presented in table 4. All CTSSs showed good interrater reliability as ICC= 0.764-0.837. CTSS2 and CTSS7 showed the largest values, (0.837 and 0.834, respectively).

AUC for ROC curves for discriminating patients in moderate from severe/critical group at the time of admission as well as related inference, threshold, sensitivity and specificity for each CTSS is presented in table 5 (upper set). Only three CTSSs namely CTSS1, CTSS2 and CTSS4 showed sufficient AUCs to be useful in triage (AUC=0.70). The sum of sensitivity and specificity for the best threshold values were 131-132% for the mentioned CTSSs. Corresponding ROC curves are shown in figure 2. Pairwise comparison of AUCs of these ROC curves by z test showed that there is no significant difference between them.

ROC curves AUCs for predicting severe/critical disease at the time of peak disease severity as well as related inference, threshold, sensitivity and specificity for each CTSS is presented in table 5 (middle set). All CTSSs showed acceptable AUCs (0.76-0.78). The sum of sensitivity and specificity for the best thresholds was 140-146% for different CTSSs. Corresponding ROC curves are shown in figure 3. Pairwise comparison of AUCs of these ROC curves showed that there is no significant difference between them.

AUC for ROC curves for predicting critical disease at the time of peak disease severity as well as related inference, threshold, sensitivity and specificity for each CTSS is also presented in table 5 (lower set). All CTSSs showed acceptable AUCs (0.77-0.79). The sum of sensitivity and specificity for the best thresholds for such diagnosis was 141-146% for different CTSSs. Corresponding ROC curves are shown in figure 4. Pairwise comparison of AUCs of these ROC curves showed that there is significant difference only in CTSS1-CTSS5, CTSS4-CTSS5, CTSS1-CTSS7 and CTSS4-CTSS7 pairs (p-value=0.04 for all four pairs) and no significant difference was present in the other pairs.

Discussion

Many researchers have used CTSS as a disease quantifying tool in COVID19 [1-10]. Some of them evaluated CTSS by ROC curve AUC, sensitivity, specificity and other indices of test performance and also by interrater reliability [6,10]. To our knowledge 6 types of CTSS have been proposed and we propose another one. We evaluated 7 CTSS types for their performance in triage and prognostication and also interrater reliability.

Because RT-PCR rarely if ever had been ordered for patients with mild symptoms in our institution, due to lack of resources, our cohort is composed of more severely affected patients in comparison with the other studies [3-10] and mortality rate was much higher (42%). As most of other mentioned studies [3-5, 8-10], men were more frequent in our cohort than women (57 vs 39). This may indicate that women are affected less, probably because of estrogen protective effect [21] or possibly they less frequently seek medical assistance in the area.

Our results showed good interrater reliability between two radiologists for all CTSSs (ICC= 0.764-0.837). the best ICCs were for CTSS2 and CTSS7, the two requiring more numerous segmentations. We failed to reproduce the brilliant interrater reliability reported in the earlier studies as ICC for CTSS3 had been reported to be 0.976 [6], but we computed 0.764. ICC for CTSS7 had been reported 0.936 [10], but we computed 0.834. The difference between previously reported ICC values and our reported ICCs can be due to two reasons. First, overall, more severe disease in our cohort, making scoring process more complex and second and more important is that we decided to compute ICCs based on 2-way random model, single measurement form and absolute agreement type ($ICC_{1,1}$ with absolute agreement) which produces the lowest ICC values, but is the most reliable one among the 10 ICC classes if reproducibility of the test is to be evaluated [17,18]. For CTSS3 the authors did not mention that what model, form and type of ICC they were reporting [6], therefore, generalization to a larger community of radiologists is not possible. The same is true for reported CTSS7 [10].

We evaluated discriminatory performance of CTSSs between the two moderate and severe/critical groups for triage. Calculated AUCs ranged 0.67-0.7 and there were only three CTSSs with sufficient ROC curve AUCs to be suitable for clinical implementation in triage of the patients, although they showed borderline value (0.70). They were CTSS1, CTSS2 and CTSS4 and their performance were far from ideal. Again, these results are not compatible with earlier studies; as for CTSS3 the reported AUC for diagnosing severe-critical disease was 0.918 (95% CI 0.962–0.985) and CTSS3 cut-off of 7.5 had 82.6% sensitivity and 100% specificity in diagnosing severe/critical group [6]. Our computed AUC value is 0.69 for AUC which is regarded as unsatisfactory. The same is true for CTSS7 with reported AUC of 0.892 (95% CI 0.814, 0.944) and that CTSS7 cut-off value of 19.5 had 83.3% sensitivity and 94% specificity in diagnosing severe/critical groups [10], but our calculated AUC is 0.67 (CI 0.56-0.78), again unsatisfactory. This discrepancy in results is most probably because of relative low incidence of severe/critical disease in the mentioned studies as their cohort included only about 10% severe/critical disease patients in CTSS3 study [6] and less than 18% in CTSS7 study [10], but in our study the corresponding percentage is 57%. We do not favor a very powerful role for CTSS in triage of patients, although some role still exists, more specifically for CTSS1, CTSS2 and CTSS4.

CTSSs performed better in prognostication than triage with acceptable AUCs for all the CTSSs both in discriminating moderate from severe/critical group and discriminating moderate/severe from critical group at peak disease severity, as all the related AUCs were acceptable for clinical use with AUCs of 0.76-0.79.

More recent reports show results compatible with our study as Hajiahmadi and colleagues reported ROC curve AUC 0.764 for CTSS1 for predicting severe/critical disease in a cohort including 24% severe/critical disease patients [22] while our calculated figure was 0.79. In addition, Aminzadeh and co-workers used a CTSS method similar to our CTSS5 and reported ROC curve AUC of 0.65 for triage of severe/critical patients and 0.76 for predicting critical disease at peak disease severity [23] and our corresponding calculated values for CTSS5 were 0.69 and 0.77 respectively.

Two limitations should be considered, one is the absence of mildly diseased patients in our cohort which was because RT-PCR was not ordered routinely for mildly diseased patients who are not hospitalized. The other one was the absence of long-term follow-up after discharge to evaluate the relation of CTSSs to long-term

sequelae of COVID-19.

Declarations

- Ethics approval and consent to participate: Our institutional review board waived the requirement to obtain written informed consent for this retrospective study, which evaluated de-identified data and involved no potential risk for patients. To avert any potential breach of confidentiality, no link between the patients and the researchers was made available.
- Consent for publication: not applicable
- Availability of data and material: The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.
- Competing interests: The authors declare that they have no competing interests.
- Funding: Iran University of Medical Sciences.
- Authors' contributions: AAN contributed to study design, data collection, statistical analysis, and manuscript writing. ASh, MAA, FD, and MB contributed to data collection, and AS contributed to study design and manuscript writing.
- Acknowledgments: We appreciate all the assistance and guidance provided by the Clinical Research Development Unit (CRDU) of Firoozabadi Hospital, without which completion of this work would not be possible.

Refrences

1. Jin YH, Cai L, Cheng ZS, Cheng H, Deng T, Fan YP, Fang C, Huang D, Huang LQ, Huang Q, Han Y. A rapid advice guideline for the diagnosis and treatment of 2019 novel coronavirus (2019-nCoV) infected pneumonia (standard version). Military Medical Research. 2020 Dec;7(1):1-23. doi: 10.1186/s40779-020-0233-6
2. Xie X, Zhong Z, Zhao W, Zheng C, Wang F, Liu J. Chest CT for typical coronavirus disease 2019 (COVID-19) pneumonia: relationship to negative RT-PCR testing. Radiology. 2020 Aug;296(2):E41-5. doi: 10.1148/radiol.2020200343
3. Zhao W, Zhong Z, Xie X, Yu Q, Liu J. Relation between chest CT findings and clinical conditions of coronavirus disease (COVID-19) pneumonia: a multicenter study. American Journal of Roentgenology. 2020 May;214(5):1072-7. doi: 10.2214/AJR.20.22976. Epub 2020 Mar 3.
4. Zhou S, Wang Y, Zhu T, Xia L. CT features of coronavirus disease 2019 (COVID-19) pneumonia in 62 patients in Wuhan, China. American Journal of Roentgenology. 2020 Jun;214(6):1287-94. doi: 10.2214/AJR.20.22975.
5. Chung M, Bernheim A, Mei X, Zhang N, Huang M, Zeng X, Cui J, Xu W, Yang Y, Fayad ZA, Jacobi A. CT imaging features of 2019 novel coronavirus (2019-nCoV). Radiology. 2020 Apr;295(1):202-7. doi: 10.1148/radiol.2020200230.
6. 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). European radiology. 2020 Aug;30(8):4407-16. https://doi.org/10.1007/s00330-020-06817-6
7. Pan F, Ye T, Sun P, Gui S, Liang B, Li L, Zheng D, Wang J, Hesketh RL, Yang L, Zheng C. Time course of lung changes on chest CT during recovery from 2019 novel coronavirus (COVID-19) pneumonia. Radiology. 2020 Feb 13. doi: 10.1148/radiol.2020200370.
8. Saeed GA, Gaba W, Shah A, Al Helali AA, Raidullah E, Al Ali AB, Elghazali M, Ahmed DY, Al Kaabi SG, Almazrouei S. Correlation between Chest CT Severity Scores and the Clinical Parameters of Adult Patients with COVID-19 pneumonia. Radiology Research and Practice. 2020 Jan 1;2021. doi: 10.1155/2021/6697677
9. Xiong Y, Sun D, Liu Y, Fan Y, Zhao L, Li X, Zhu W. Clinical and high-resolution CT features of the COVID-19 infection: comparison of the initial and follow-up changes. Investigative radiology. 2020. doi: 10.1097/RLI.0000000000000674
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. Radiology: Cardiothoracic Imaging. 2020

Mar 30;2(2):e200047. https://doi.org/10.1148/ryct.2020200047
11. van der Jagt EJ, Smits HJ. Cardiac size in the supine chestfilm. European journal of radiology. 1992 May 1;14(3):173-7. doi:https://doi.org/10.1016/0720-048X(92)90080-S
12. Kim SJ, Im JG, Kim IO, Cho ST, Cha SH, Park KS, Kim DY. Normal bronchial and pulmonary arterial diameters measured by thin section CT. Journal of computer assisted tomography. 1995 May 1;19(3):365-9. doi: 10.1097/00004728-199505000-00005
13. Matsuoka S, Uchiyama K, Shima H, Ueno N, Oish S, Nojiri Y. Bronchoarterial ratio and bronchial wall thickness on high-resolution CT in asymptomatic subjects: correlation with age and smoking. American Journal of Roentgenology. 2003 Feb;180(2):513-8. doi: 10.2214/ajr.180.2.1800513.
14. Islamic Republic of Iran, Ministry of Health and Medical Education, [Guide to the diagnosis and treatment of Covid-19 disease at the levels of outpatient and inpatient services] https://web.ssu.ac.ir/Dorsapax/userfiles/Sub28/09.pdf
15. Ajlan AM, Ahyad RA, Jamjoom LG, Alharthy A, Madani TA. Middle East respiratory syndrome coronavirus (MERS-CoV) infection: chest CT findings. American journal of roentgenology. 2014 Oct;203(4):782-7. doi: 10.2214/AJR.14.13021.
16. Pandey RM. Commonly used t-tests in medical research. Journal of the Practice of Cardiovascular Sciences. 2015 May 1;1(2):185. doi:10.4103/2395-5414.166321
17. Trevethan R. Intraclass correlation coefficients: clearing the air, extending some cautions, and making some requests. Health Services and Outcomes Research Methodology. 2017 Jun 1;17(2):127-43. doi: 10.1007/s10742-016-0156-6
18. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med. 2016; 15 (2): 155–63. doi: 10.1016/j.jcm.2016.02.012
19. David W.. Hosmer, Lemeshow S, Rodney X.. Sturdivant. Assessing the fit of the model. In: Applied logistic regression. 2nd edition. New York: Wiley; 2000. Pp. 160–164.
20. Hajian-Tilaki K. Receiver operating characteristic (ROC) curve analysis for medical diagnostic test evaluation. Caspian journal of internal medicine. 2013;4(2):627-635.
21. Dangis A, De Brucker N, Heremans A, Gillis M, Frans J, Demeyere A, Symons R. Impact of gender on extent of lung injury in COVID-19. Clinical Radiology. 2020 Jul 1;75(7):554-6. doi: 10.1016/j.crad.2020.04.005
22. Hajiahmadi S, Shayganfar A, Janghorbani M, Esfahani MM, Mahnam M, Bakhtiarvand N, Sami R, Khademi N, Dehghani M. Chest Computed Tomography Severity Score to Predict Adverse Outcomes of Patients with COVID-19. Infection & chemotherapy. 2021 Jun;53(2):308.
23. Aminzadeh B, Layegh P, Foroughian M, Tavassoli A, Emadzadeh M, Teimouri A, Maftouh M. Evaluation of the Prognostic Value of Chest Computed Tomography Scan in COVID-19 Patients. Iranian Journal of Radiology. 2021 Apr 30;18(2).

Table 1- clinical severity of COVID19

Measured Indicator/Severity [a]	Mild	Moderate	Severe	Critical
Respiratory Rate	[?]24	[?]30	-	-
SPO_2	[?]93	93>SPO_2[?]90	89>SPO_2[?]85	<85 [b]
Respiratory Distress	None	None	Mild to moderate	Severe [c]
Blood Pressure	-	-	-	<90/60

a: presence of any of the severity indicators of the more severe group places the patient in the more severe group

b: despite high-flow O_2 administration

c: nasal flaring , air hunger, intercostal retraction, subcostal retraction

Table 2: Seven proposed COVID-19 CT severity score systems

CTSSs	Segmentation	Severity Score for each segment	Maximum Score
CTSS1 [2,3]	Three zones in each lung are divided by carina and lower pulmonary vein	1-4 according to percentage of involvement (<25, 25-49, 50-74, >75)	24
CTSS2 [4]	The same zonal concept as CTSS1 with additional division of each zone into anterior and posterior regions divided by midpoint of diaphragm antero-posteriorly	1-4 according to percentage of involvement (<25, 25-49, 50-74, >75)	48
CTSS3 [5,6]	Five anatomic lobes of the lungs	1-4 according to percentage of involvement (<25, 25-49, 50-74, >75)	20
CTSS4 [7,8]	Five anatomic lobes of the lungs	1-5 according to percentage of involvement (>5, 5-25, 25-49, 50-74, >75)	25
CTSS5 [current authors]	Five anatomic lobes of the lungs with additional consideration of the lingula as a separate lobe	1-5 according to percentage of involvement (>5, 5-25, 25-49, 50-74, >75)	30
CTSS6 [9]	Five anatomic lobes of the lungs	1-4 according to the diameter of the largest lesion in each lobe (<1cm, 1-3cm, >3cm up to 50% of the lobe, >50% of a lobe	20
CTSS7 [10]	18 anatomic segments of the lung with an additional division of apico-posterior segment of the left upper lobe into apical and posterior divisions and anteromedial segment of the left lower lobe into anterior and medial segments	No involvement=0 <50% involvement=1 [?]50% involvement=2	40

Table 3- patients demographics and distribution of disease severity at presentation and at peak disease severity

	Number (Male/Female)	Mean age ± SD
Total	96 (57/39)	63.6 ± 17.4
Moderate disease at presentation	41 (25/16)	57.3 ± 18.9
Severe disease at presentation	53 (31/22)	68.2 ± 14.9
Critical disease at presentation	2 (1/1)	71.5 ± 6.4
Moderate disease at peak severity	22 (13/9)	52.5 ± 20.1
Severe disease at peak severity	31 (17/14)	62 ± 16.8
Critical disease at peak severity	43 (27/16)	70.4 ± 12.9
Discharged	56 (32/24)	59 ± 18.6
Deceased	40 (25/15)	70 ± 13.3

Table 4- interrater reliability between the two radiologists and related inference

CT Severity Score	Intraclass Correlation	Inference about Interrater Reliability
CTSS1	0.783	good
CTSS2	0.837	good
CTSS3	0.764	good
CTSS4	0.778	good
CTSS5	0.784	good
CTSS6	0.773	good
CTSS7	0.834	good

Table 5- AUC, confidence interval, related inference, best threshold and related sensitivity and specificity for ROC curves about different CTSSs about diagnosis of severe/critical group at presentation and at peak disease severity and also for diagnosis of critical disease at peak severity

	Average CTSS	AUC for ROC Curve	95% Confidence Inte
diagnosis of severe/critical patients at presentation	CTSS1	0.70	0.59-0.80
	CTSS2	0.70	0.60-0.81
	CTSS3	0.69	0.58-0.80
	CTSS4	0.70	0.59-0.80
	CTSS5	0.69	0.58-0.80
	CTSS6	0.68	0.57-0.79
	CTSS7	0.67	0.56-0.78
diagnosis of severe/critical patients at peak disease severity	CTSS1	0.78	0.67-0.88
	CTSS2	0.78	0.68-0.89
	CTSS3	0.76	0.65-0.87
	CTSS4	0.77	0.66-0.88
	CTSS5	0.77	0.65-0.88
	CTSS6	0.76	0.65-0.87
	CTSS7	0.77	0.65-0.88
diagnosis of critical patients at peak disease severity	CTSS1	0.79	0.70-0.88
	CTSS2	0.78	0.69-0.87
	CTSS3	0.78	0.69-0.87
	CTSS4	0.79	0.70-0.88
	CTSS5	0.77	0.68-0.86
	CTSS6	0.76	0.67-0.86

	Average CTSS	AUC for ROC Curve	95% Confidence Inte
	CTSS7	0.79	0.70-0.88

Figure 1- flowchart for patient selection

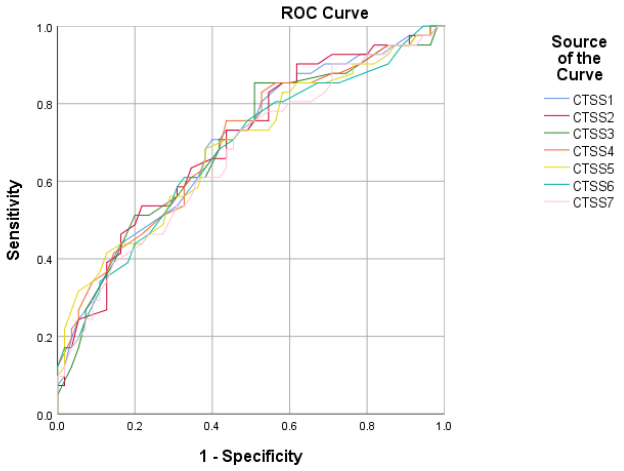

Figure 2- ROC curves plotted for different average CTSSs discriminating moderate from severe/critical disease at the time of hospital admission

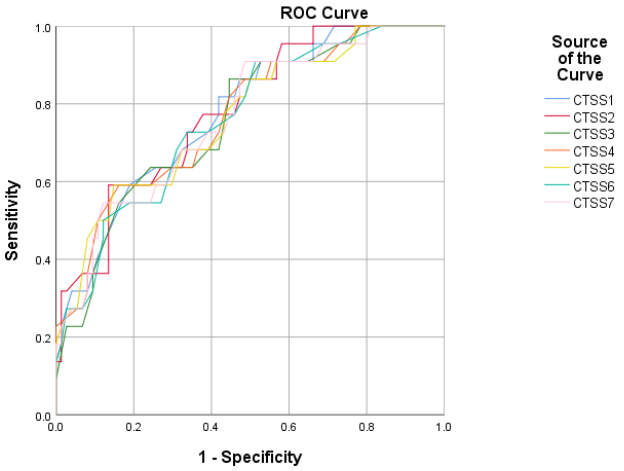

Figure 3- ROC curves plotted for different average CTSSs discriminating moderate from severe/critical disease in the most severe disease period

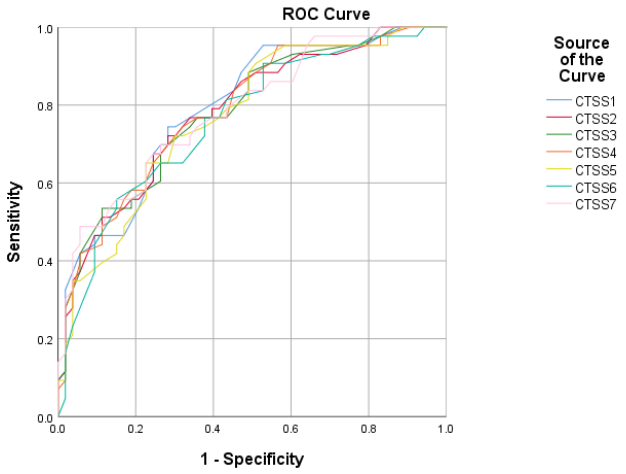

Figure4- ROC curves plotted for different average CTSSs discriminating moderate/severe from critical disease in the most severe disease period