The value of AI based CT severity scoring system in triage of patients with Covid-19 pneumonia as regards oxygen requirement and place of admission

Anirudh Kohli, Tanya Jha¹, Amal Babu Pazhayattil
Departments of Imaging and 'Critical Care, Breach Candy Hospital Trust, Breach Candy, Cumballa Hill, Mumbai, Maharashtra, India

Correspondence: Dr. Tanya Jha, Breach Candy Hospital Trust, 60A, Bhulabhai Desai Road, Breach Candy, Cumballa Hill, Mumbai, - 400 026, Maharashtra, India. E-mail: tanyajha@gmail.com

Abstract

Context: CT scan is a quick and effective method to triage patients in the Covid-19 pandemic to prevent the healthcare facilities from getting overwhelmed. Aims: To find whether an initial HRCT chest can help triage patients by determining their oxygen requirement, place of treatment, laboratory parameters and risk of mortality and to compare 3 CT scoring systems (0-20, 0-25 and percentage of involved lung models) to find if one is a better predictor of prognosis than the other. Settings and Design: This was a prospective observational study conducted at a Tertiary care hospital in Mumbai. Patients undergoing CT scan were included by complete enumeration method. Methods and Material: Data collected included demographics, days from swab positivity to CT scan, comorbidities, place of treatment, laboratory parameters, oxygen requirement and mortality. We divided the patients into mild, moderate and severe based on 3 criteria - 20 point CT score (OS1), 25 point CT score (OS2) and opacity percentage (OP). CT scans were analysed using CT pneumonia analysis prototype software (Siemens Healthcare version 2.5.2, Erlangen, Germany). Statistical Analysis: ROC curve and Youden’s index were used to determine cut off points. Multinomial logistic regression used to study the relations with oxygen requirement and place of admission. Hosmer-Lemeshow test was done to test the goodness of fit of our models. Results: A total of 740 patients were included in our study. All the 3 scoring systems showed a significant positive correlation with oxygen requirement, place of admission and death. Based on ROC analysis a score of 4 for OS1, 9 for OS2 and 12.7% for OP was determined as the cut off for oxygen requirement. Conclusions: CT severity scoring using an automated deep learning software programme is a boon for determining oxygen requirement and triage. As the score increases, the chances of requirement of higher oxygen and intubation increase. All the three scoring systems are predictive of oxygen requirement.

Key words: Covid-19; HRCT chest; oxygen requirement

Introduction

SARS CoV 2 infection, which initially began in Wuhan, China in December 2019 has been declared as a pandemic...
The main concern of a pandemic is the overwhelming of medical facilities due to the sheer number of patients presenting at the same time. This may also trigger panic in the general population leading to societal issues. Lockdowns are primarily to slow the number of patients presenting at the same time, allowing medical facilities to try and cope. An important component of alleviating the pressure on medical establishments is to triage patients. COVID-19 pneumonia has a wide spectrum of presentations and outcomes, ranging from asymptomatic to severe hypoxia which may result in death. Triage is needed to decide the order of treatment when the number of patients is large, outweighing the capacity of the available facilities. As the clinical presentation, management, and outcomes of COVID-19 is so varied, triaging patients into those who require management at home, covid care facilities or those who require hospitalization is essential. Additionally, for the patients in hospital, whether they will require admission in a ward or ICU, whether the oxygen requirement would be low flow, high flow oxygen, or whether they would require non-invasive ventilation or mechanical ventilation via intubation is also important. The role of CT for triaging patients has been emerging.

COVID-19 pneumonia pathologically is a diffuse alveolar damage, which progresses in a temporal manner. Initially patients may be mildly symptomatic but can rapidly progress to severe hypoxia requiring oxygen support and hospitalisation. The extent of lung involvement in the second phase of diffuse alveolar damage approximately between day 7 and 10 may be a good surrogate to decide disease burden. A low percentage of aerated lung tissue correlates with a poor prognosis. For this purpose, a CT severity score has been proposed to determine the extent of lung involvement based on extent of involvement on CT scan.

To test this hypothesis, we undertook a study to determine whether extent of lung involvement on a CT scan can help triage. The main components of which include the requirement for oxygen, type of oxygen support required as well as location of admission in the health care facility.

**Subjects and Methods**

This was a prospective observational study conducted at Breach Candy Hospital Trust, which is a tertiary care hospital at Mumbai. Laboratory, clinical, and radiological data of all patients who underwent Ct scan for covid between April and September 2020 was collected. Ethical committee waived the need for ethical clearance and informed consent for this non interventional study.

All patients with RT PCR positive for SARS COV2 who underwent Ct scan in our hospital were included. All CT scans were done within one week of disease onset. We excluded patients with 2 consecutive negative RT PCR, patients lost to follow up, and patients with underlying chronic lung disease like ILD, organising pneumonias, carcinoma lung and Tb sequelae.

We collected the baseline demographics – age, gender and comorbidities. Laboratory data collected included IL6, D-dimer, CRP, procalcitonin and ferritin levels. The blood tests were done within 48 hours of CT scan. We also noted the need for hospitalisation, need for ICU admission, medical therapy administered and mortality. Of note, we differentiated between the oxygen requirement of patients as those on room air, with low oxygen requirement (up to 6 L/min), high oxygen requirement (7-15 L/min), NIV or HFNC, and need for intubation and mechanical ventilation. All patients undergoing CT scan from August 2020 to October 2020 were included by complete enumeration method and were followed up for as long as they were admitted in the hospital, resulting either in death or discharge.

We evaluated all the patients based on 3 criteria- 20 point CT score (OS1), 25 point CT score (OS2) and opacity percentage (OP). For each of the 3 criteria, patients were divided into mild moderate and severe group based on the cut offs as shown in Table 1.

**CT analysis**

All chest CT scans were performed using a 16 slice CT scanner (Siemens Biograph Horizon) with the following parameters: 130 kV, tube voltage 100-200 mAs, rotation time 0.6 s, pitch 1:35. 1 mm slice thickness, sharp convolution kernel reconstructions with a window width of 1200 HU and a window length -600 HU was performed. The scanning range was from the apex of the lung to the costophrenic angle. No IV contrast was administered. The scan was captured in the end-inspiratory phase, whenever it was possible for the patient to hold the breath adequately. Using CT pneumonia analysis prototype software (Siemens Healthcare version 2.5.2, Erlangen, Germany), an AI algorithm based on three-dimensional segmentation automatically. The software automatically detected and quantified abnormal tomographic patterns (ground-glass opacities and consolidations) in each and both lung parenchyma based on deep learning and deep reinforcement learning.

The CT was assessed for features of COVID-19 infection, like ground-glass opacities, consolidation, crazy paving pattern,
Results

After exclusion, 740 patients were included in our study. Majority of them belonged to the mild group (64%, 69% and 73% based on OS1, OS2 and OP, respectively). This was followed by moderate group and the least number of patients were in the severe group (based on OS1 27% and 9%, OS2 22% and 9% and on OP 16% and 11% belonged to moderate and severe group respectively). Thus, the distribution of the patients was not equal between the 3 groups and the patient characteristics between the 3 categories also varied.

The mean age of the patients was 59 years and there were 65% males (n = 483). The mean age was lesser in the mild group as compared to moderate and severe. The mean duration between swab positivity and CT was 2 days. 34·18% of the patients had no comorbidities. Of the patients with comorbidities, the most common ones were diabetes mellitus in 41·58%, hypertension in 44·6% and ischemic heart disease in 15·4%. The remaining comorbidities along with their distribution between the 3 groups and the treatment given are as shown in Table 3.

Relation with oxygen requirement

80·75% of the patients were on room air. Majority of them belonged to the mild group (77%, 81% and 85% based on OS1, OS2 and OP respectively). Only 0·67% from OS, 0·67% from OS2 and 1·5% from OP patients from severe group were on room air. 10% of the patients were on low oxygen (oxygen requirement 1-6 L/min). Majority of them belonged to mild group followed by moderate group. Only 22% (based on OS1), 21% (based on OS2) and 29% (based on OP) of the patients were from severe group. A clear majority of patients who were on high oxygen, HFNC/NIV and intubated belonged to the severe group as shown in Table 4.

We ran a multinomial logistic regression to study the relation of OS1, OS2 and OP with oxygen requirement. The regression shows a significant positive relation between the scores and oxygen requirement for all 3 categories (p-value = 0·000). The models show that OS1 and OS2 were slightly better than OP to predict the oxygen requirement. The adjusted prediction graphs, Figures 2-4, show how the probability of a patient in
room air (Outcome = 0) decreases as OS1, OS2 and OP increase, while the probability of a patient being intubated (Outcome = 4) increases as OS1, OS2 and OP increase. ROC curve analysis showed a score of 4 for OS1 (AUC-0·9230, Youden Index – 0·668304 Sensitivity – 61% specificity – 98%), 9 for OS2 (AUC-0·9197 Youden index-0·6711637)
Kohli, et al.: CT severity scoring system for triage in patients with Covid-19 pneumonia

We ran a multinomial logistic regression to study the relation of OS1, OS2 and OP with place of admission. The regression showed that the odds of a patient being admitted to Ward and ICU significantly increases with an increase in OS1, OS2 and OP (p value = 0.000). Based on the pseudo r-squared values, model using OS1 performed marginally better than the models using OS2 and OP. The adjusted prediction graphs, Figures 5-7, show the probability of a patient being admitted to OPD decreases and ICU increases as the scores increase.

Relation with laboratory parameters

The laboratory parameters assessed included D-dimer, Ferritin, Interleukin 6, Procalcitonin and CRP. The values were available for only 390, 285, 268, 54 and 188 patients respectively as the lab tests were done depending upon the clinical condition of the patient and on discretion of the treating consultant. The median values of the lab parameters for each group are mentioned in Table 6. Even though the values are higher as the severity increases, no statistically significant correlation could be found between lab values and severity scores. The correlation matrix between laboratory values and Ct scores is shown in Figure 8.

Relation with mortality

There were 23 deaths in our study group. The distribution is as shown in the bar graph [Figure 9]. Majority of deaths were in the severe group. The relationship between severity score and mortality was found to be statistically significant in all the 3 groups (P value= 0.000). The correlation between mortality and severity scores is as shown in Figures 10-12.

Discussion

A number of CT severity scoring systems have been proposed. There are 20, 25, 40 and 72 point scales which evaluate extent of lung involvement depending upon the involvement of each lobe, expressed as percentage, which is finally summed up to provide the final score.\[^{8-11}\] This may be done subjectively with visual interpretation of the scans or by an automated deep learning software programme.\[^{12,13}\] The subjective visual method is fraught with numerous inter and intra observer errors as it is a visual assumption, resulting in significant under and over estimation. This process is manual; thus, it is extremely time consuming.

sensitivity- 58% specificity – 97%) and 12.7% according to OP (AUC= 0.9135 Youden Index= 0.6644 sensitivity – 58% specificity – 97%) as the cut off for oxygen requirement.

Relation with place of admission

Majority of the patients were admitted in ward (51%) followed by OPD (44%). Almost all the patients treated on OPD basis belonged to the mild group. Only 4%, 2% and 05% patients from moderate group in OS1 OS2 and OP category were treated as OPD patients. Majority of the patients from severe category were admitted in the ICU (74% in OS1, 76% in OS2 and 83% in OP group) and the remaining were admitted in ward. Majority of patients with moderate disease were admitted in ward (73% in OS1, 72% in OS2 and 60% in OP group). The distribution of patients is as shown in Table 5.
Kohli, et al.: CT severity scoring system for triage in patients with Covid-19 pneumonia

which does not help in a pandemic where accurate and timely information is required for triage. In view of this we utilized an automated deep learning software - Siemens Healthcare version 2.5.2, Erlangen, Germany.

The extent of lung involvement was correlated with oxygen requirement, laboratory parameters, place of admission and mortality.

We found a statistically significant relation between the severity scores and oxygen requirement, need for hospital and ICU admission and mortality. There was no significant relation between the scores and laboratory parameters, which could probably be because laboratory parameters were available for a limited number of patients in our study. Since only the early CT scans, done within one week of disease, were taken into consideration, these findings can help in early triaging of the patients who are likely to deteriorate and administer early medical therapy even if the patients clinically have mild symptom.

Since a number of scoring systems are available, we compared the 20, 25 point scales and percentage of lung involvement to determine whether one predictor is better than other. Our study found that all three scoring systems related well with the oxygen requirement, though the performance of 0-20 scale model was slightly better. Thus, any of the scoring systems are equally predictive of patient outcome and can be used for assessing the prognosis of patients. On ROC curve analysis we found a cut off of 4 for OS1, 9 for OS2 and 12.5% for OP to be a predictor of oxygen requirement.

Majority of the patients in the mild group were on room air and were treated on an OPD basis whereas majority of the patients in severe group were admitted in ICU and were either on NIV/HFNC or were intubated. This finding was consistent for all the 3 scoring systems. Mortality was highest in the severe group. Unlike the findings of Zhang et al., we could not find a correlation between the scoring system and the laboratory parameters.

The review of relevant literature is as shown in Table 7.
This was a prospective study conducted with an appropriate sample size. This study demonstrated a direct relationship between the CT severity scores and oxygen requirement and need for intubation, independent of the laboratory values and comorbidities of the patients. To the best of our knowledge, no study has provided a comparison between the different scoring systems for CT severity. This study showed that all 3 scoring systems were predictive and none was significantly superior to the other. While most of the studies provided a cut off for predicting mechanical ventilation and oxygen requirement based on opacity...
Optimal CT severity score threshold for identifying deceased patients was a percentage of lung involvement less than 73% being a predictor of ICU admission or death.

A cut-off of 23% of lung involvement showed distinguished critically ill patients from patients with less severe disease.

Threshold for 51% for mortality and 25% for mechanical ventilation.

As the severity increased, the number of lobe involved and CT severity score increased from 4 to 5 and 6 to 12 respectively. A cut off of 5 helped to identify cases with severe pneumonia (i.e. SpO2 less than 93% on room air and P/F < 300).

The cut-off value of total CT scores was determined to be 16.5 for CT score ≥ 13 was related to poor 5-day outcome.

The 25-point CT severity score correlates well with the Covid-19 clinical severity.

Optimal CT severity score threshold for identifying deceased patients was 10. The mean score of survivors was 7 and deceased patients was 14.

In conclusion, an early CT scan in patients affected with Covid-19 is predictive of the oxygen requirement of the patient. As severity scores increase the chances of requirement of higher oxygen and intubation increase. The severity scoring system can be based on lobar involvement and scored 0-20 or 0-25 or based on percentage of lung involved, as they are all predictive of oxygen requirement. CT severity scoring using an automated deep learning software programme is a great boon for determining oxygen requirement and triage.

Acknowledgements
Would like to acknowledge the contribution of Sanya Jha for data analysis and interpretation in the study.

Financial support and sponsorship
Nil.

Conflicts of interest
There are no conflicts of interest.
References