ORIGINAL ARTICLE

Independent Role of CT Chest Scan in COVID-19 Prognosis: Evidence From the Machine Learning Classification

Sudhir Bhandari,[1] Ajit Singh Shaktawat,[1] Amit Tak,[2] Bhoopendra Patel,[3] Jitentdra Gupta,[4] Kapil Gupta,[4] Shivankan Kakkar,[5] Kshitij Darshan Shah,[6] Ayesha Arora,[7] Amitabh Dube[4]

(1) Department of Medicine, SMS Medical College and Hospitals, Jaipur, Rajasthan, India.
(2) Department of Physiology, RVRS Medical College and Hospitals, Bhilwara, Rajasthan, India.
(3) Department of Physiology, All India Institute of Medical Sciences, Bilaspur, Himachal Pradesh, India.
(4) Department of Physiology, SMS Medical College and Hospitals, Jaipur, Rajasthan, India.
(5) Department of Pharmacology, SMS Medical College and Hospitals, Jaipur, Rajasthan, India.
(6) Arthrocare Hospital, Ahmedabad, Gujarat, India.
(7) Chicago School of Professional Psychology, Chicago, IL, USA.

Abstract

Background: The current coronavirus disease-19 (COVID-19) pandemic call attention to the key role informatics play in healthcare. The present study discovers an independent role of computerised tomography chest (CT) scans in prognosis of COVID-19 using classification learning algorithms.
Methods: In this retrospective study, 57 RT PCR positive COVID-19 patients were enrolled from SMS Medical College, Jaipur (Rajasthan, India) after approval from the Institutional Ethics Committee. A set of 21 features including clinical findings and laboratory parameters and chest CT severity score were recorded. The CT score with mild, moderate and severe categories was chosen as response variable. The dimensionality reduction of feature space was performed and classifiers including, decision trees, K-nearest neighbours, support vector machine and ensemble learning were trained with principal components. The model with highest accuracy and area under the ROC curve (AUC) was selected.
Results: The median age of patients was 55 years (range: 20-99 years) with 37 males. The feature space was reduced from 21 to 7 predictors, that included fever, cough, fibrin degradation products, haemoglobin, neutrophil-lymphocyte ratio, ferritin and procalcitonin. The linear support vector machine was chosen as the best classifier with 73.7 % and 0.69 accuracy and AUC for severe CT chest score, respectively. The variance contributed by first three principal components were 97.5 %, 2.4 % and 0.0 %, respectively.
Conclusion: In view of low degree of relationships between predictors and chest CT scan severity score category as interpreted from accuracy and AUC it can be concluded that chest CT scan has an independent role in the prognosis of COVID-19 patients.

Key words: Classification; COVID-19; Chest CT scan; Machine learning; Pandemic.

Correspondence:
AMIT TAK
E: dramittak@gmail.com

ARTICLE INFO
Received: 17 October 2021
Revision received: 23 November 2021
Accepted: 23 November 2021

Introduction

Globally, there have been 76,023,488 confirmed cases and 1,694,128 deaths due to COVID-19 as on 22 December 2020.[1] A large number of mathematical models and machine learning methods have been proposed to characterise the various aspects of pandemic.[2] The machine learning algorithms have been used in automated analysis of CT images during COVID-19 pandemic to reduce the effort of clinicians.[3-5] The World Health Organization recommends RT PCR test for the diagnosis of COVID-19.[1] The low sensitivity of RT PCR and high sensitivity of chest computerised tomography (CT) for COVID-19 make chest CT a useful tool in the diagnosis.[6] A number of studies showed

Copyright © 2021 Bhandari et al. This is an open access article distributed under the Creative Commons Attribution License (CC BY), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
This article should be cited as follows: Bhandari S, Shaktawat AS, Tak A, Patel B, Gupta J, Gupta K, et al. Independent role of CT chest scan in COVID-19 prognosis: evidence from the machine learning classification. Scr Med 2021 Sep;52(4):273-8.

relationship between disease severity and chest CT severity score in COVID-19. Similarly, studies showed relationship of clinical and laboratory parameters with disease severity.

In addition to usefulness of CT chest in diagnosis, its key role in prognostication and further management of COVID-19 patients has been proposed. Furthermore, CT scan is important in the patient triage and from logistics point of view. The objective of the present study was to evaluate independent role of CT chest in the prognosis of COVID-19 patients utilising various machine learning classifiers.

Methods

A retrospective cross-sectional study was planned to discover independent role of CT chest in COVID-19 prognosis. A total of 57 RT-PCR SARS CoV-2 positive patients were enrolled from SMS Medical College, Jaipur (Rajasthan, India) after obtaining approval from the institutional Ethics Committee. The clinical findings and laboratory parameters along with CT chest score of COVID-19 patients were recorded. The relationship between CT score category with clinical and laboratory parameters was evaluated using classifiers while selecting CT score category as response variable. The CT chest has definite role in prognosis of COVID-19, but this study was undertaken to assess the independent role of CT chest in prognosis of COVID-19 as compared to clinical and laboratory parameters.

Feature selection

A set of 21 features included age, sex, history of diabetes mellitus, history of hypertension, complaints of fever, cough, shortness of breath, sore throat, haemoglobin (Hb) in grams %, total leukocyte count (TLC) in 1,000 cells per mm^3, platelet count (PC) in 100,000 per mm^3, Differential Neutrophil Count (DNC) in percent, Differential Lymphocyte Count (DLC) in percent, neutrophil-lymphocyte count ratio (NLR), fibrin degradation products (FDP) in mg/mL, D-dimer in ng/mL FEU, activated partial thromboplastin time (APTT) in seconds, prothrombin time (PT) in seconds, international normalised ratio (INR), procalcitonin in ng/mL and ferritin in ng/mL. The CT scoring was based on lobe involvement, as suggested by Li et al.[14] A score was assigned for each lobe on the basis of its involvement: score 0 for 0 % involvement, score 1 for less than 5 % involvement, score 2 for 5 % to 25 % involvement, score 3 for 26 % to 49 % involvement, score 4 for 50 % to 75 % involvement and score 5 for greater than 75 % involvement. There was a score of 0 to 5 for each lobe, with a total possible score lie between 0 to 25.[7] The CT score was converted from ordinal to categorical response variable so machine learning classification can be applied. The CT score from 0 to 11 was considered as mild, from 12 to 17 as moderate and from 18 to 25 as severe.

Dimensionality reduction

The features and response variable were imported in the classification leaner app.[8] The classifiers were trained with individual predictors and accuracy and area under the ROC curve (AUC) of each model was recorded. Among them, 10 most accurate predictors were selected. Correlation analysis and parallel coordinate plot (PCP) was used for further feature selection and visualisation. A total of 7 features were selected from 10 features (Figure 1). To further reduce the dimensionality of feature space, principal component analysis (PCA) was performed. In view of 57 observations, considering 10 observations per feature, 5 features were used to train the classifiers with 5-fold cross validation.[9]

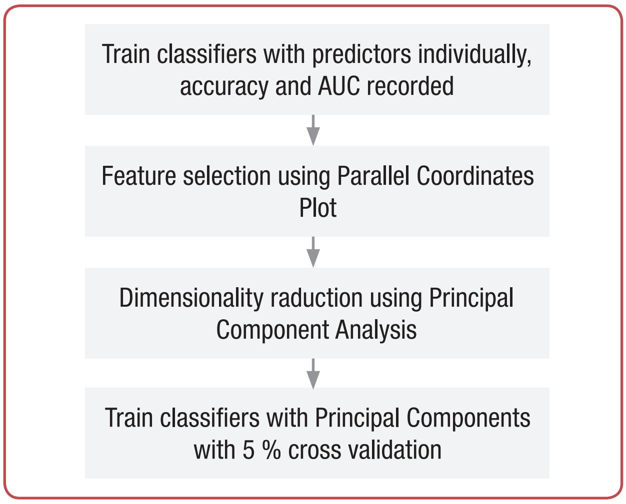

Figure 1: Flowchart showing steps performed to reduce dimensionality of feature space

Classifiers and training

The classifiers used in the application include decision trees, support vector machine (SVM), K nearest neighbours and ensemble learning classifiers. These classifiers used different methodology for classification.[10] The decision trees include complex tree, medium tree and simple tree classifiers. The SVM include linear SVM, quadratic SVM, cubic SVM, fine Gaussian SVM, medium Gaussian SVM and coarse Gaussian SVM classifiers. The ensemble classifiers include boosted trees, bagged trees and RUS boosted trees classifiers.

Statistical analysis

The accuracy and AUC for various classification models were calculated using classification learner app in MATLAB 2016a (MATLAB Team, 2020).[9] A low degree of relationship between CT score and predictors can be interpreted as more likelihood of independent role of CT scan and vice-versa.

Results

The individual features (number of features = 21) of 57 patients (median: 55 years, range 20-99 years; 37 males) with COVID-19 were trained with CT score as response variable. The top ten predictors with highest accuracy were selected. These include NLR, fever, PT, procalcitonin, cough, FDP, Hb, DNC, DLC and ferritin (Table 1). Correlation analysis showed that PT was highly correlated with FDP ($r = 0.78$; $p < 0.001$); DNC was highly correlated with NLR ($r = 0.85$; $p < 0.001$) and, DLC ($r = - 0.85$; $p < 0.001$) showed high correlation with NLR and DNC ($r = - 0.98$; $p < 0.001$). From the above features PT, DNC and DLC were removed. The relationships can be visualised in parallel coordinate plots (Figure 2 and 3). Thus, two qualitative features, fever and cough and five quantitative features including FDP, Hb, NLR, ferritin and procalcitonin were selected. As there were 2 categorical features (as PCA is not applicable to categorical variables), 3 principal components were chosen manually to get a total of 5

Table 1: The decision tree, support vector machines (SMV), K-nearest neighbour (KNN) and ensemble classifiers were trained with individual features. The classifier model with highest accuracy and area under the curve (AUC) for each feature is shown

Variable	Classifier	AUC	Accuracy (%)
NLR	Ensemble Bagged Trees	0.87	75.40
Fever	Fine KNN and Ensemble Subspace KNN	0.50	64.90
PT	Fine Gaussian SVM	0.80	64.90
Procalcitonin	Fine Gaussian SVM	0.74	64.90
Cough	Fine KNN and Ensemble Subspace KNN	0.50	63.20
FDP	Medium Gaussian SVM	0.64	63.20
	Ensemble Boosted Trees	0.58	63.20
DNC	Complex and Medium Tree	0.61	63.20
DLC	Ensemble Bagged Trees	0.70	63.20
Ferritin	Complex and Medium Tree	0.69	63.20
Age	Cubic and Medium KNN	0.63	61.40
APPT	Simple Tree	0.68	61.40
INR	Quadratic SVM	0.70	61.40
TLC	Fine Gaussian SVM	0.47	61.40
Shortness of Breath	Ensemble RUS Boosted Trees	0.72	59.60
PC	Fine KNN	0.60	59.60
Sex	Cubic SVM	0.56	57.90
Sore Throat	Cubic SVM	0.70	57.90
History of Hypertension	Coarse Gaussian SVM	0.63	57.90
History of Diabetes Mellitus	Cubic SVM	0.51	57.90
d-Dimer	Coarse KNN	0.44	57.90

PCT: procalcitonin; FDP: fibrin degradation products; PT: prothrombin time; Hb: haemoglobin; DNC: differential neutrophil count; NLR: neutrophil-lymphocyte ratio; DLC: differential lymphocyte count.

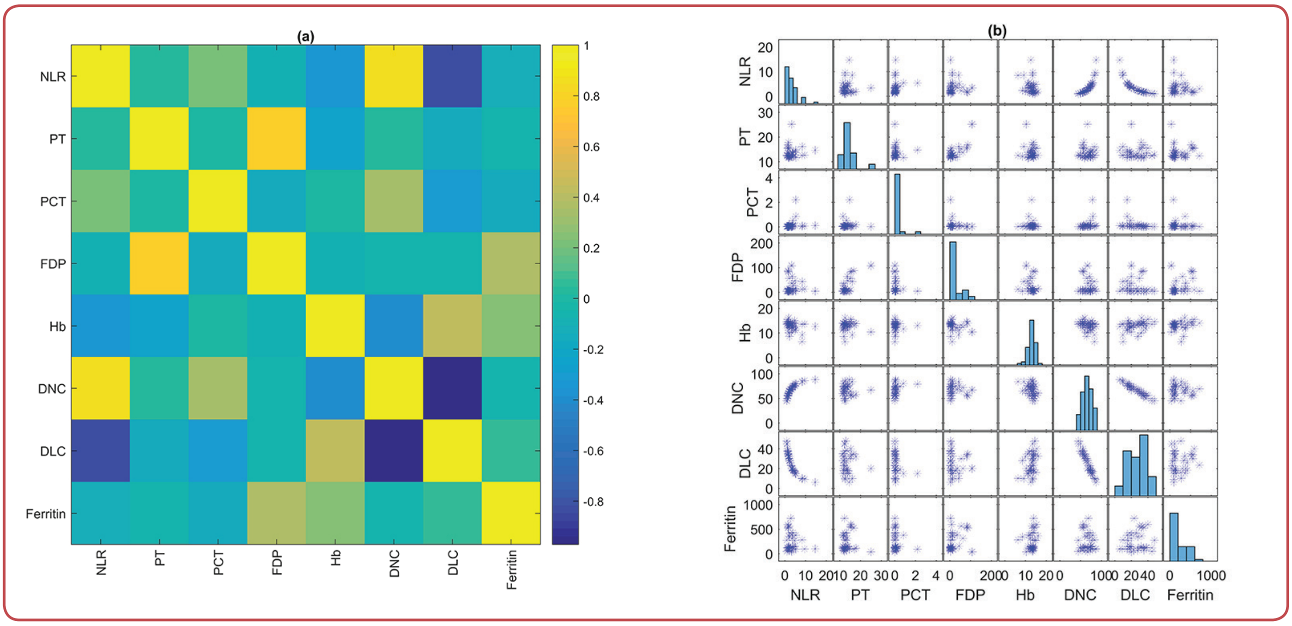

Figure 2: Correlation analysis (a) Heat map shows positive correlation between FDP and PT; DNC and NLR; negative correlation between DLC and NLR; DLC and DNC (b) scatter map shows correlation among selected features.
PCT: procalcitonin; FDP: fibrin degradation products; PT: prothrombin time; Hb: haemoglobin; DNC: differential neutrophil count; NLR: neutrophil-lymphocyte ratio; DLC: differential lymphocyte count.

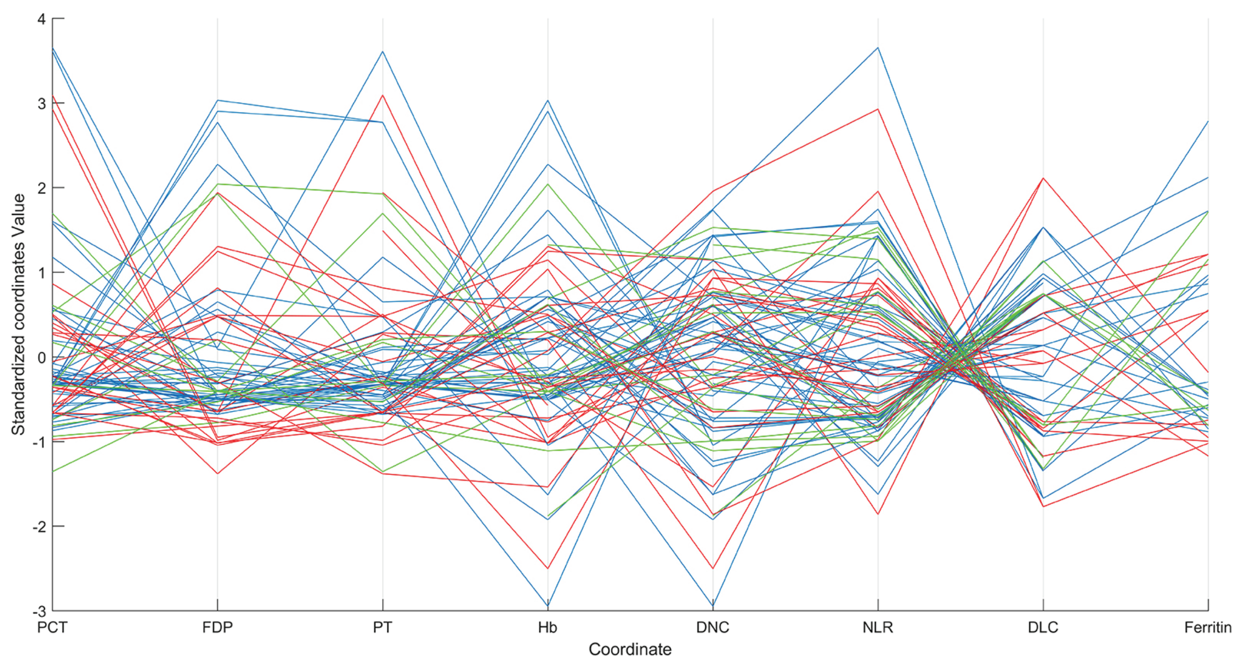

Figure 3: Shows parallel coordinate plots to visualise relationship among various features. NLR and DLC features has shown inverse relationship. PCT: procalcitonin; FDP: fibrin degradation products; PT: prothrombin time; Hb: haemoglobin; DNC: differential neutrophil count; NLR: neutrophil-lymphocyte ratio; DLC: differential lymphocyte count.

features to train the models. The best model was linear support vector machine with 73.7 % and 0.69 accuracy and AUC for severe CT chest level respectively. The variance contributed by three principal components were 97.5 %, 2.4 % and 0.0 %, respectively.

Discussion

As per World Health Organization, a confirmed case of COVID-19 is defined as a patient with RT-PCR test–positive for SARS CoV-2, irrespective of clinical signs and symptoms. The pooled estimate of sensitivity of RT-PCR tests for SARS CoV-2 is 89 % (95 % CI: 81 - 94 %). Thus, one or more negative results do not rule out COVID-19.[11] To compensate the shortcomings of the low sensitivity and time-consuming process of RT-PCR, chest CT examination has an auxiliary and key role in the diagnosis and subsequent management of COVID-19 patients.

First, CT chest is more sensitive than RT-PCR test as concluded from various studies. A meta-analysis included six studies comprising a total of 1431 patients who were mainly symptomatic and at high risk for COVID-19, reported a chest CT pooled sensitivity of 94.6 % (95 % CI: 91.9 - 96.4 %) and a pooled specificity of 46.0 % (95 % CI: 31.9 - 60.7 %) in the detection of COVID-19.[6, 12] Similarly, Caruso showed sensitivity, specificity and accuracy of CT were 97 % (95 % CI: 88 - 99 %), 56 % (95 % CI: 45 - 66 %) and 72 % (95 % CI: 64 - 78 %), respectively.[13] The misdiagnosis rate of CT chest scan to diagnose COVID-19 is quite low when taking reverse transcriptase polymerase chain reaction as gold standard (3.9 %).[14] The CT chest scan has a key role in diagnosis of COVID-19.[15, 16] In more than 70 % of RT-PCR test positive cases of SARS CoV-2, CT chest findings include ground-glass opacities, vascular enlargement, bilateral abnormalities, lower lobe involvement and posterior predilection. However, about 10 % to 70 % patients revealed consolidation (51.5 %), linear opacity (40.7 %), septal thickening and/ or reticulation (49.6 %), crazy-paving pattern (34.9 %), air bronchogram (40.2 %), pleural thickening (34.7 %), halo sign (34.5 %), bronchiectasis (24.2 %), nodules (19.8 %), bronchial wall thickening (14.3 %) and reversed halo sign (11.1 %). The uncommon findings include pleural effusion (5.2 %), lymphadenopathy (5.1 %), tree-in-bud sign (4.1 %), central lesion distribution (3.6 %), pericardial effusion (2.7 %) and cavitating lung lesions (0.7 %).[17]

Second, CT chest can be used as a follow-up tool to monitor the disease evolution and evaluate the severity of COVID-19 patients for its invasiveness and objectivity. Furthermore, CT can predict the prognosis.[18] The present study emphasises the independent role of CT in prognosis of COVID-19 patients. The use of a chest CT severity score may be useful for standardised assessment of the degree of pulmonary involvement in COVID-19 for prognostication purposes.[19] Based on the CT chest score, disease severity can be classified into mild, moderate and severe. Though fever (80 %, 74 - 87 %) and cough (53 %, 33 - 72 %) are prevalent symptoms in COVID-19, there was no correlation with disease severity as evaluated in the present study.[20] In case of laboratory parameters, except NLR no other parameter was significantly related to CT chest score. However, Xiong showed relationship between CT severity and C-reactive protein, erythrocyte sedimentation rate and lactate dehydrogenase.[21] In his study, Durhan found no relationship between COVID-19 disease severity and fever ($p = 0.82$), dry cough ($p = 0.46$), diabetes ($p = 0.60$), hypertension ($p = 0.29$), haemoglobin ($p = 0.92$). Though he showed significant correlation between CT severity score and age ($p = 0.001$), sex ($p = 0.002$), neutrophil count ($p = 0.03$), lymphocyte count ($p = 0.01$), NLR ($p = 0.004$), platelet count ($p = 0.03$), ferritin ($p < 0.001$), CRP ($p < 0.001$), procalcitonin ($p < 0.001$) and D-dimer ($p < 0.001$).[22] Hana et al found platelet count is not associated with COVID-19 severity.[23] In a retrospective study of 313 patients, ferritin and CRP levels were significantly higher in patients with severe CT findings compared to the patients with mild and moderate CT findings ($p < 0.05$).[24] The chest CT score had positive associations with total leukocyte count, CRP, ESR, procalcitonin and a negative association with lymphocyte count.[25] In a similar study, correlation analysis showed that the CT chest score was significantly correlated with lymphocyte count, monocyte count, C-reactive protein, procalcitonin, days from illness onset and body temperature ($p < 0.05$).[26] In view of low degree of relationships between predictors and CT score category it can be concluded that chest CT scan has independent role in prognosis of COVID-19.

In addition to above, the Fleischner Society recommends CT imaging: (a) to establish a baseline pulmonary status; (b) to facilitate risk stratification in patients with comorbidities and (c) in patients with moderate to severe symptoms of COVID-19.[27] The current COVID-19 pandemic has highlighted the essential role of chest CT examination in patient triage in the emergency departments, allowing them to be referred to "COVID" or "non-COVID" wards.[28] Rubin et al emphasised the use of CT chest in diagnosis of COVID-19 from logistic point of view.[27]

Conclusion

Despite inconsistencies in relationships between CT chest score and other parameters, most of the relationships are not clinically significant and an independent role of CT chest in the prognosis and further management of COVID-19 patients is proposed. Furthermore, due to the low sensitivity of RT PCR, CT chest scan is recommended for disease management.

Limitations of the study

The sample size of the study is low. The study includes the routine clinical and laboratory parameters, however, independence of CT chest scan with other laboratory parameters needs to be tested.

Ethical Statement

The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by Ethical Committee of the SMS Medical College, Jaipur (No 429 dated 2020 Jun 26) and the individual consent for this retrospective analysis was waived.

Acknowledgements

The contributors to the article acknowledge the Principal, SMS Medical College and Attached Hospitals, Jaipur (Rajasthan, India) and Department of Medicine for availability of data and their ongoing support against the menace of the global pandemic of COVID-19.

Conflict of interest

None.

References

1. World Health Organization [Internet]. WHO Coronavirus Disease (COVID-19) Dashboard. Geneva. [Cited 2020-Dec-22]. Available from: https://covid19.who.int/.
2. Tak A, Das B, Shah M, Dia S, Dia M, Gahlot S. COVID-19 and lockdown in India: evaluation using analysis of covariance. J Antivir Antiretrovir 2021;13(3):216.
3. Kang H, Xia L, Yan F, Wan Z, Shi F, Yuan H, et al. Diagnosis of coronavirus disease 2019 (COVID-19) with structured latent multi-view representation learning. IEEE Trans Med Imaging 2020 Aug;39(8):2606-14.
4. Ardakani AA, Kanafi AR, Acharya UR, Khadem N, Mohammadi A. Application of deep learning technique to manage COVID-19 in routine clinical practice using CT images: Results of 10 convolutional neural networks. Comput Biol Med 2020 Jun;121:103795. doi: 10.1016/j.compbiomed.2020.103795.
5. Ni Q, Sun ZY, Qi L, Chen W, Yang Y, Wang L, et al. A deep learning approach to characterize 2019 coronavirus disease (COVID-19) pneumonia in chest CT images. Eur Radiol 2020 Dec;30(12):6517-27.
6. Adams HJA, Kwee TC, Yakar D, Hope MD, Kwee RM. Systematic review and meta-analysis on the value of chest CT in the diagnosis of coronavirus disease (COVID-19): Sol Scientiae, Illustra Nos. AJR Am J Roentgenol 2020 Dec;215(6):1342-50.
7. Li K, Wu J, Wu F, Guo D, Chen L, Fang Z, et al. The clinical and chest CT features associated with severe and critical COVID-19 pneumonia. Invest Radiol 2020 Jun;55(6):327-31.
8. Indrayan A, Malhotra RK. Medical Biostatistics. 4th ed. Florida, USA: In Relationships: Qualitative Dependent, CRC Press, Taylor & Francis Group; 2018. p. 472-3.
9. MATLAB Team, Statistics and Machine Learning Toolbox 10.2, Classification Learner App, MATLAB. Version 9.0.0.341360 (R 2016a). Natick, Massachusetts: The Mathworks Inc; R2015.
10. Stuart J. Russell and Peter Norvig. Artificial Intelligence: A Modern Approach, Global Edition. Artificial Intelligence: A Modern Approach, Global Edition; 2016.
11. World Health Organization, 2020 [Internet]. Laboratory testing for 2019 novel coronavirus (2019-nCoV) in suspected human cases. [Cited 2020-Dec-22]. Available from: https://www.who.int/publications-detail/laboratorytesting-for-2019-novel-coronavirus-in-suspected-humancases-20200117.
12. Raptis CA, Hammer MM, Short RG, Shah A, Bhalla S, Bierhals AJ, et al. Chest CT and coronavirus disease (COVID-19): a critical review of the literature to date. AJR Am J Roentgenol 2020 Oct;215(4):839-42.
13. Caruso D, Zerunian M, Polici M, Pucciarelli F, Polidori T, Rucci C, et al. Chest CT features of COVID-19 in Rome, Italy. Radiology 2020 Aug;296(2):E79-E85.
14. Li Y, Xia L. Coronavirus disease 2019 (COVID-19): Role of chest CT in diagnosis and management. AJR Am J Roentgenol 2020 Jun;214(6):1280-6.
15. Tenda ED, Yulianti M, Asaf MM, Yunus RE, Septiyanti W, Wulani V, et al. The importance of chest CT scan in COVID-19. Acta Med Indones 2020 Jan;52(1):68-73.
16. Aljondi R, Alghamdi S. Diagnostic value of imaging modalities for COVID-19: Scoping review. J Med Internet Res 2020 Aug 19;22(8):e19673. doi: 10.2196/19673.
17. Kwee TC, Kwee RM. Chest CT in COVID-19: What the radiologist needs to know. RadioGraphics 2020;40(7):1848–65.
18. Li M. Chest CT features and their role in COVID-19. Radiol Infect Dis 2020 Jun;7(2):51-4.
19. 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 Mar 30;2(2):e200047. doi: 10.1148/ryct.2020200047.
20. Wan S, Li M, Ye Z, Yang C, Cai Q, Duan S, et al. CT manifestations and clinical characteristics of 1115 patients with coronavirus disease 2019 (COVID-19): a systematic review and meta-analysis. Acad Radiol 2020 Jul;27(7):910-21.
21. Xiong Y, Sun D, Liu Y, Fan Y, Zhao L, Li X, et al. Clinical and high-resolution CT features of the COVID-19 infection: comparison of the initial and follow-up changes. Invest Radiol 2020 Jun;55(6):332-9.
22. Durhan G, Düzgün SA, Demirkazık FB, Irmak İ, İdilman İ, Akpınar MG, et al. Visual and software-based quantitative chest ct assessment of COVID-19: correlation with clinical findings. Diagn Interv Radiol 2020 Nov;26(6):557-64.
23. Hana C, Aboulenain S, Dewaswala N, Narendran V. Does thrombocytopenia truly correlate with COVID-19 severity? Blood 2020 Nov 5;136:39–40.
24. Altinbilek E, Ozturk D, Ozkurt H, Oncul A, Ocal M, Tumer M, et al. Comparison of laboratory results and chest CT findings in COVID-19 patients. Ann Clin Analytical Med 2020:587-91.
25. Zhang J, Meng G, Li W, Shi B, Dong H, Su Z, et al. Relationship of chest CT score with clinical characteristics of 108 patients hospitalized with COVID-19 in Wuhan, China. Respir Res 2020 Jul 14;21(1):180. doi: 10.1186/s12931-020-01440-x.
26. Wu J, Wu X, Zeng W, Guo D, Fang Z, Chen L, et al. Chest CT findings in patients with coronavirus disease 2019 and its relationship with clinical features. Invest Radiol 2020 May;55(5):257-61.
27. Rubin GD, Ryerson CJ, Haramati LB, Sverzellati N, Kanne JP, Raoof S, et al. The role of chest imaging in patient management during the COVID-19 pandemic: a multinational consensus statement from the Fleischner Society. Chest 2020 Jul;158(1):106-16.
28. Jalaber C, Lapotre T, Morcet-Delattre T, Ribet F, Jouneau S, Lederlin M. Chest CT in COVID-19 pneumonia: A review of current knowledge. Diagn Interv Imaging 2020 Jul-Aug;101(7-8):431-7.