Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Quantification of COVID-19 Opacities on Chest CT – Evaluation of a Fully Automatic AI-approach to Noninvasively Differentiate Critical Versus Noncritical Patients

Christoph Mader, MD, Simon Bernatz, MD, Sabine Michalik, Vitali Koch, MD, Simon S. Martin, MD, Scherwin Mahmoudi, MD, Lajos Basten, MD, Leon D. Grünewald, MD, Andreas Bucher, MD, Moritz H. Albrecht, MD, Thomas J. Vogl, MD, Christian Booz, MD

Objectives: To evaluate the potential of a fully automatic artificial intelligence (AI)-driven computed tomography (CT) software prototype to quantify severity of COVID-19 infection on chest CT in relationship with clinical and laboratory data.

Methods: We retrospectively analyzed 50 patients with laboratory confirmed COVID-19 infection who had received chest CT between March and July 2020. Pulmonary opacifications were automatically evaluated by an AI-driven software and correlated with clinical and laboratory parameters using Spearman-Rho and linear regression analysis. We divided the patients into sub cohorts with or without necessity of intensive care unit (ICU) treatment. Sub cohort differences were evaluated employing Wilcoxon-Mann-Whitney-Test.

Results: We included 50 CT examinations (mean age, 57.24 years), of whom 24 (48%) had an ICU stay. Extent of COVID-19 like opacities on chest CT showed correlations (all $p < 0.001$ if not otherwise stated) with occurrence of ICU stay (R = 0.74), length of ICU stay (R = 0.81), lethal outcome (R = 0.56) and length of hospital stay (R = 0.33, $p < 0.05$). The opacities extent was correlated with laboratory parameters: neutrophil count (NEU) (R = 0.60), lactate dehydrogenase (LDH) (R = 0.60), troponin (TNTHS) (R = 0.55) and c-reactive protein (CRP) (R = 0.51). Differences ($p < 0.001$) between ICU group and non-ICU group concerned longer length of hospital stay (24.04 vs. 10.92 days), higher opacity score (12.50 vs. 4.96) and severity of laboratory data changes such as c-reactive protein (11.64 vs. 5.07 mg/dl, $p < 0.01$).

Conclusions: Automatically AI-driven quantification of opacities on chest CT correlates with laboratory and clinical data in patients with confirmed COVID-19 infection and may serve as non-invasive predictive marker for clinical course of COVID-19.

Key Words: COVID-19; SARS-CoV-2 infection; Pneumonia; Viral; Artificial Intelligence; Chest-CT.

Abbreviations: AI Artificial intelligence, **ARDS** Acute respiratory distress syndrome, **BIL** Bilirubin, **COVID-19** Coronavirus disease 2019, **CRP** C-reactive protein, **CT** Computed tomography, **DDI** D-dimers, **DI2IN** Deep Image to Image Network, **DICOM** Digital Imaging and Communications in Medicine, **GGO** Ground-glass opacities, **GTP** Alanine aminotransferase, **HR-CT** High-resolution computer tomography, **HST** Urea, **HU** Hounsfield units, **ICU** Intensive care unit, **IL-6** Interleukin-6, **KREA** Creatinine, **LAC** Lactate, **LDH** Lactate dehydrogenase, **LEU** White blood cell count, **LYM** Lymphocyte count, **NEU** Neutrophil count, **PACS** Picture archiving and communication system, **PCT** Procalcitonin, **PHO** Percentage of high opacity, **RIS** Radiology information system, **RT-PCR** Real-time reverse transcription polymerase-chain-reaction, **SARS-CoV-2** Severe acute respiratory syndrome coronavirus type, **THR** Thrombocyte count, **TNTHS** Troponin, **TPZ** Quick value, **VRT** Volume rendering technique, **WHO** World Health Organization

Acad Radiol 2021; 28:1048–1057

From the Department of Diagnostic and Interventional Radiology, University Hospital Frankfurt, Frankfurt am Main, Germany (C.M., S.B., S.M., V.K., S.S.M., S.M., L.B., L.D.G., A.B., M.H.A., T.J.V., C.B.); Dr. Senckenberg Institute for Pathology, University Hospital Frankfurt, Frankfurt am Main, Germany (S.B.). Received January 6, 2021; revised February 14, 2021; accepted March 1, 2021. **Address to correspondence to:** C.M. e-mail: Christoph.mader@kgu.de

https://doi.org/10.1016/j.acra.2021.03.001

INTRODUCTION

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) can cause coronavirus disease 2019 (COVID-19) which was first reported in Wuhan, China in December 2019 (1). The virus has spread worldwide at enormous speed and burdens especially old and multimorbid patients as well as healthcare systems. Globally, as of 13

February 2021, there have been 107.686.655 confirmed cases of COVID-19, including 2.368.571 deaths, reported to WHO (2). The disease frequently starts with flu-like symptoms like fever, dry cough, fatigue and can lead to acute respiratory distress syndrome, organ failure and intensive care unit (ICU) admission (approximately 30%) with consecutive high mortality rates (up to 15%) (3). Although real-time reverse transcription polymerase-chain-reaction (RT-PCR) is considered to be the current gold standard for diagnosing, the method is slow and has a high false negative rate (4). Among the 1014 COVID-Patients in Wuhan up to February 6, 2020 only 59% had a positive RT-PCR result, whereas 88% showed positive findings on chest computed tomography (CT) scans (4). Chest CT therefore may not only be able to detect more cases of viral pneumonia in terms of COVID-19 but also to quantify the severity of pulmonary COVID-19 manifestation (5). Frequently reported findings in chest CT were bilateral and peripheral ground-glass opacities (GGO), crazy-paving pattern, and consolidations (6).

Some studies have analyzed the relationship between the severity of abnormalities on chest CT scans in COVID-19 positive patients with clinical and laboratory data. Xiong et al. showed that laboratory data like CRP, erythrocyte sedimentation rate and LDH correlated with the severity of abnormalities manually quantified on chest CT scans (7). Lanza et al. found that semi-automatic quantification of compromised lung volume in COVID-19 patients on chest CT scans correlated with the need for oxygenation support and intubation and might therefore serve as a triaging and prediction tool in COVID-19 positive patients (8). However, manual and semi-automatic quantification and feature extraction of pulmonary opacifications is time-consuming and prone to inter-observer variance complicating implementation in clinical routine (9). In this context, the application of fully automatic deep learning models may enable radiologists to make a more accurate diagnosis and quantification of corresponding lung changes in terms of severity assessment (9-12). In addition, software-based analysis of chest CT scans may improve time-efficiency in clinical routine, which gets more and more important in context of increasing examination numbers. Anyhow, there has been no diagnostic model based on medical imaging without high risk of bias so far. Studies often lack information regarding how regions of interest are annotated which results in no transparency and reproducibility of the models (13).

Recently, a new software prototype for fully automatic quantification of COVID-19 like lung opacities on chest CT scans has been developed. This prototype offers several advantages in that it is fast, fully-automatic, provides two combined severity measures of the disease and can be used on CT scans with a slice thickness up to 5 mm (14).

In this initial study, we aimed at evaluating the potential of this prototype for fully automatic quantification of COVID-19 like opacities on chest CT scans of 50 laboratory confirmed COVID-19 patients in relation with clinical and laboratory data.

MATERIALS AND METHODS

Ethics Approval

The study was approved by the ethical review board of our institution. Written informed consent was obtained for experimentation with human subjects. All investigations were conducted in accordance with the 1964 Declaration of Helsinki and its amendments.

Patient selection and study design

This retrospective study was approved by the local ethics committee. Written informed consent was waived. Two radiologists (⋆BLINDED⋆ and ⋆BLINDED⋆) scanned our Picture Archiving and Communication System (PACS) and Radiology Information System (RIS) for a total of 151 patients, who had received chest CT scans at our institution in case of suspicion for a COVID-19 infection in the period from March 1, 2020 to July 31, 2020. COVID-19 infection was RT-PCR-approved in 50 cases with a positive real-time fluorescence polymerase chain reaction (RT-PCR) assay for SARS-CoV-2 nucleic acid with nasopharyngeal or oropharyngeal swab specimens. Negative SARS-CoV-2 RT-PCR as well as patients who did not fulfill inclusion criteria (children <18 years, slice thickness >5 mm, artifacts) were excluded (Fig 1). Demographics, typical COVID-19 symptoms, and laboratory data were retrospectively collected from the electronic medical records platform. CRP, LDH, TNTHS, NEU, procalcitonin (PCT), d-dimers (DDI), white blood cell count (LEU), lymphocyte count (LYM), thrombocyte count (THR), interleukin-6 (IL-6), bilirubin (BIL), lactate (LAC), alanine aminotransferase (GTP), quick value (TPZ), creatinine (KREA) and urea (HST) were collected at the time point of CT performance (±48 h) and were analyzed to examine a possible relation with automated quantification of abnormalities associated with COVID-19. Time interval between onset of symptoms and imaging at our institution, length of hospital stays, length of ICU stay and outcome of the patients were collected. For sub cohort analysis, patients were divided into a critical ($n = 24$) and non-critical group ($n = 26$) based on necessity of an ICU stay.

CT Image Data Acquisition

Unenhanced chest CT examinations were carried out on two 192-detector row, thirdgeneration CT scanners (SOMATOM Force, Siemens Healthineers, Forchheim, Germany (critical patients, $n = 10$; non-critical, $n = 12$) and SOMATOM Definition AS, Siemens Healthineers [critical, $n = 14$; non-critical, $n = 14$]). A tube voltage of 120 kV and automatic tube current modulation (100−400 mA) were used. Images were reconstructed with a slice thickness between 1.00 mm to 5.00 mm (1.00 mm = 42/50 [0.84], 5.00 mm = 8/50 [0.16]). The average time interval between initial chest CT performance and onset of symptoms was 13.08 days (range, 2−45 days).

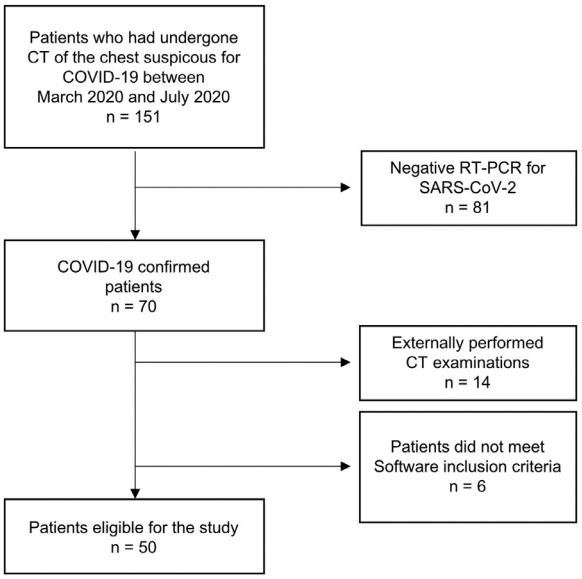

Figure 1. Flow chart of patient inclusion.

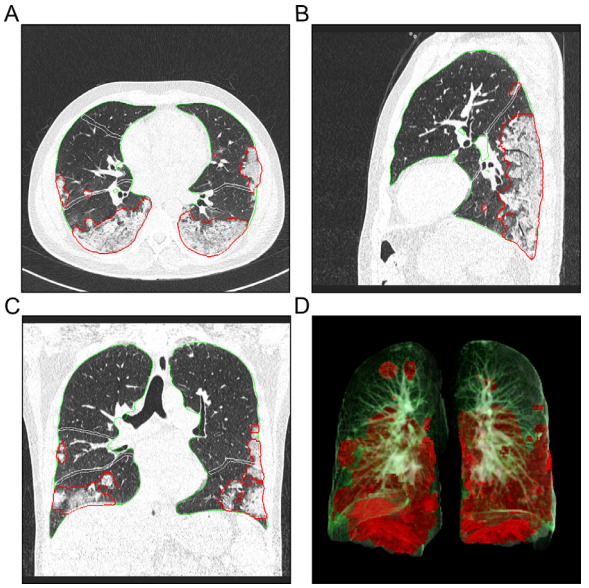

Figure 2. (a-c) CT images showing postprocessing analysis and quantification results of the evaluated AI-driven software prototype. Segmentation of the lung lobes is shown in different colors. Segmentation of lung opacities is displayed in red. Images illustrate a 1.0 chest HR-CT of a 35-year-old male patient with confirmed COVID-19 infection. CT series show typical COVID-19 peripherally pronounced GGO and consolidations. The CT examination was taken 5 days after symptoms of fever, dyspnea and dry cough started. The patient had an ICU stay for 9 days and could then be transferred back to the normal infirmary. He had a hospital stay of 11 days in total. Currently the patient is largely symptom-free, apart from a discreet tiredness and an intermittent slight pulling in the lungs. (d) The VRT allows for quick overview of the spatial distribution of opacities. GGO, ground-glass opacities; VRT, volume rendering technique. (Color version of figure is available online.)

Software Prototype

The software used to quantify COVID-19 like abnormalities of the lung calculates percentage of opacity (PO) and lung severity score (LSS) by segmenting three-dimensional abnormalities like ground glass opacities and consolidations, that were related to COVID-19. Thereby PO and LSS define the expanse of lung involvement and allocation of involvement across lobes. High opacity abnormalities like consolidations and subsolid regions were measured by percentage of high opacity (PHO) and lung high opacity score, which define the extent of high opacity abnormalities and the distribution across the lung lobes. By default, a Hounsfield Unit (HU) threshold of -200 HU was applied for identification of high opacities. A lobe was classified as "affected" if the algorithm detected high opacity abnormalities in the defined lung part.

The overall extent of opacities is displayed by a reproducible, fully automatically calculated so called opacity score ranging from 0 - 20, which is a sum up of the five lung lobes. An opacity score of zero means that none of the lobes is affected whereas 20 represents extensive involvement of all five lobes. Lobe-wise, the score is calculated on the percentage of opacity per lobe, using a Likert scale from 0 to 4. Percentage of opacity - which represents the percentage of predicted volume of abnormalities compared with the total lung volume within a given region - is classified as follows: score = 0,1,2,3,4 extent of opacity within a given region $\leq$1%, $\leq$25%, $\leq$50, $\leq$75%, >75%, respectively (Fig 3).

The software is based on AI algorithms to automatically identify and quantify lung hyperdensities for research purposes (Fig 2). Multi-scale deep reinforcement learning (15) was used for anatomical landmark detection and a Deep Image to Image Network (DI2IN) (16) for segmentation, based on a dataset of 8792 CT volumes. To ensure a stable lobe-wise segmentation and identification of abnormal patterns, the DI2IN was trained on 8087 patients with various diseases and afterwards fine-tuned on another dataset of 1136 patients with similar abnormalities as COVID-19 like GGO and consolidation (e.g. viral pneumonia, interstitial lung diseases). The algorithm was finally evaluated on 100 control volumes without pathological chest-CT findings and 100 patients with confirmed COVID-19. The Pearson correlation coefficient between method prediction and ground truth for COVID-19 positive scans was calculated as 0.92 for PO ($p<$ .001), 0.97 for PHO ($p<$ .001), 0.91 for LSS ($p<$ .001), 0.9 for LHOS ($p<$ .001) (14). For further information we recommend the recently published study by Chaganti et al (14).

Automatic segmentation was visually approved by two radiologists (*BLINDED* and *BLINDED*) in each examination. In no cases pleural effusions or other findings affected the software-based severity measurement. No manual adjustment was performed in order to reduce manual manipulation and to examine a fully automated quantification approach.

Statistical Analysis

Statistical analysis was conducted using RStudio (Version 1.2.5001, RStudio, Inc.). Demographics and clinical

Total Opacity Score: 9
Percentage of opacity: 26.94

L/R Lobes

	Both lungs	Left lung	Right lung
Affected	Yes	Yes	Yes
Opacity Score	9	4	5
Lung volume (ml)	5369.72	2541.54	2828.18
Volume of opacity (ml)	1446.71	708.33	738.38
Percentage of opacity	26.94	27.87	26.11
Volume of high opacity (ml)	565.29	270.44	294.85
Percentage of high opacity	10.53	10.64	10.43
Mean HU total	-659.13	-653.99	-663.75
Mean HU of opacity	-311.51	-318.19	-305.09
Standard deviation total	328.41	332.39	324.72
Standard deviation of opacity	302.78	296.99	308.09

Figure 3. Representative software analysis key data to quantify the extent of overall and lobe-wise opacity abnormalities based on the 3D segmentations of lesions, lungs, and lobes in a 35-year-old male patient with confirmed COVID-19 infection. Total opacity score (range, 0-20) represents the mean of the opacity scores of each lung lobe. Percentage of opacity represents the percentage of opacity for the whole lung.

characteristics were analyzed applying descriptive statistics. Correlation coefficients were calculated using Spearman-rho. Differences between groups were calculated using Wilcoxon-Mann-Whitney-Test. Levels of significance are depicted as followed: $\star p < 0.05$; $\star\star p < 0.01$; $\star\star\star p < 0.001$.

RESULTS

Demographics, Clinical and Laboratory Data

The study included a total of 50 patients (44 male). Mean age was 57.24 years (range, 35−85 years). Common comorbidities of the patients included hypertension (17/50, 34%), cardiovascular disease (13/50, 26%), diabetes (10/50, 20%), chronic kidney disease (7/50, 14%). chronic lung disease (6/50, 12%) and malignancy (4/50, 8%). Mean time interval between onset of COVID-19 symptoms and CT image acquisition was 13.08 days (range, 2−45). The average length of hospital stay was 17.22 days (range, 0−70). One non-ICU patient was treated in an ambulant setting. A total of 24 patients were admitted to the ICU (ICU-group) with a mean length of ICU stay of 15.62 days ±13.46. Lethal disease progression was observed in 8 ICU patients (16%, ICU-group). Most common symptoms were dyspnea (33/50, 66%), fever (32/50, 64%), and dry cough (28/50, 56%). Further frequently encountered symptoms were fatigue, head and body aches (12/50, 24%) gastrointestinal symptoms like diarrhea (10/50, 20%) as well as anosmia and ageusia (8/50, 16%) (Table 1).

Chest CT Software Prototype Parameters

All CT scans ($n = 50/50$) showed opacities in both lungs. Overall opacity score ranged from 1 ($n = 2$) to 20 ($n = 4$). The mean of the score of both lungs was 8.58 (±5.34). The

TABLE 1. Demographic and Clinical Patient Characteristics

Clinical characteristics	*n*	%
Age, in years (mean, range)	57.24 (35-85)	
<40	6	0.12
40-68	34	0.68
>68	10	0.20
Male	44	0.88
Pre-existing comorbidities	*n*	%
Cardiovascular disease	13	0.26
Diabetes	10	0.20
Hypertension	17	0.34
Malignancy	4	0.08
Lung disease (2 COPD, 3 asthma, 1 silicosis)	6	0.12
Chronic kidney disease	7	0.14
No. of comorbidities (mean)	1.68	Min: 0, Max: 5
Symptoms (*n* = 44)	*n*	%
Fever	32	0.64
Cough	28	0.56
Dyspnea	33	0.66
Gastrointestinal symptoms	10	0.20
Fatigue, head & body aches	12	0.24
Anosmia & ageusia	8	0.16
Clinical features (*n* = 44)	*n*	%
Time interval between onset of symptoms & image acquisition (days, mean)	13.08	Min: 2, Max: 45
Length of hospital stay (days, mean)	17.22	Min: 0 Max: 70
Intensive care unit stay (ICU)	24	0.48
Length of Intensive care unit stay (days, mean)	15.62	Min: 1; Max: 70
lethal outcome	8	0.16

TABLE 2. Software Parameters

Number of Affected Lobes	*n*	
Both lungs	50	
Left lung	50	
Right lung	48	
Left upper lobe	46	
Left lower lobe	50	
Right upper lobe	44	
Right middle lobe	42	
Right lower lobe	48	
Opacity score (1-20)	Mean	SD
Both lungs	8.58	5.34
Left lung	3.52	2.07
Right lung	5.06	3.40
Left upper lobe	1.46	0.99
Left lower lobe	2.06	1.19
Right upper lobe	1.56	1.23
Right middle lobe	1.36	1.12
Right lower lobe	2.14	1.29
Percentage of opacity (%)	Mean	SD
Both lungs percentage of opacity	28.40	29.70
Left lung percentage of opacity	26.58	29.02
Right lung percentage of opacity	30.09	30.85
Left upper lobe percentage of opacity	20.23	27.26
Left lower lobe percentage of opacity	36.43	34.84
Right upper lobe percentage of opacity	25.15	31.54
Right middle lobe percentage of opacity	19.56	30.10
Right lower lobe percentage of opacity	41.10	34.58
Percentage of high opacity (%)	Mean	SD
Both lungs percentage of high opacity	10.46	14.55
Left lung percentage of high opacity	9.40	14.05
Right lung percentage of high opacity	11.38	15.21
Left upper lobe percentage of high opacity	6.30	12.44
Left lower lobe percentage of high opacity	14.83	19.74
Right upper lobe percentage of high opacity	8.28	14.43
Right middle lobe percentage of high opacity	5.23	11.50
Right lower lobe percentage of high opacity	16.46	18.95

right lung revealed higher mean opacity score of 5.06 (left lung, 3.52). The mean percentage of opacity was 28.40 (±29.70) for both lungs. The percentage of high opacity was 10.46 (±14.55) for both lungs (Table 2). The prototype had a constant average evaluation time of approximately 2 minutes for automatic quantification of a chest 1.0 HR-CT scan measured on a sample ($n = 10$).

Correlations Between Software Parameters and Clinical and Laboratory Data

We could reveal highest correlation coefficients (all $p < 0.001$ if not otherwise stated) between the opacity score and clinical data such as occurrence of an ICU stay (R = 0.74), length of ICU stay (R = 0.81), length of hospital stay (R = 0.33, $p < 0.05$), and mortality (R = 0.56). Furthermore, moderate to strong correlations between laboratory data and the opacity score were found for numerous blood compounds among others LEU (R = 0.55), NEU (R = 0.60), CRP (R = 0.51), IL-6 (R = 0.47, $p < 0.05$), LDH (R = 0.6) and TNTHS (R = 0.59) (Table 3; Figs 4 and 5).

TABLE 3. Relationship Between Opacity Score and Clinical and Laboratory Characteristics

Clinical Characteristics	Opacity Score	*N*
Age	0.20	50
Sex	0.32*	50
Dyspnea	0.37**	50
Fever	-0.36	50
Cough	-0.11	50
Fatigue	-0.16	50
Anosmia & aneugesia	-0.06	50
Gastrointestinal symptoms	-0.07	50
Cardiovascular disease	0.25	50
Chronic kidney disease	-0.01	50
Chronic lung disease	0.03	50
History of cancer	-0.15	50
Hypertension	0.00	50
Diabetes	0.12	50
No. of comorbidities	0.09	50
Intensive care unit	0.74***	50
Intensive care unit length of stay	0.81***	50
Length of stay	0.33*	50
Lethal outcome	0.56***	50
White blood cell count/nl	0.55***	50
Neutrophil count/nl	0.60***	48
Lymphocyte count/nl	-0.14	50
Thrombocyte count/nl	0.1	50
Neutrophil to lymphocyte ratio	0.61***	48
C-reactive protein mg/dl	0.51***	50
Procalcitonin ng/ml	0.56***	40
D-dimers ng/ml	0.32*	44
Troponin pg/ml	0.55***	47
Lactate dehydrogenase U/L	0.60***	50
Interleukin-6 pg/ml	0.47**	44
Lactate mg/dl	0.46**	39
Creatinine mg/dl	0.20	50
Urea mg/dl	0.51***	43
Bilirubin mg/dl	0.30*	50
Alanine aminotransferase U/L	0.41**	50
Quick value%	-0.13	48

Data are R correlation coefficients, calculated using Spearman correlation.

* Correlation is significant at the 0.05 level.

** Correlation is significant at the 0.01 level.

*** Correlation is significant at the 0.001 level.

Sub cohort-analysis of ICU and non-critical (non-ICU) patients showed significant differences (all $p < 0.001$ if not otherwise stated) with ICU patients revealing longer hospital stay (24.04 vs. 10.92 days), higher opacity score (12.50 vs. 4.96), higher percentage of opacity (50.06 vs 8.42) and higher PHO (19.53 vs. 2.09) (Fig 6). Non-invasive quantification value differences for ICU and non-ICU patient sub cohorts were corroborated by significant differences in laboratory data (all $p < 0.001$ if not otherwise stated) such as LEU (8.45 vs. 5.20), NEU (7.01 vs. 3.38), TNTHS (84.91 vs.

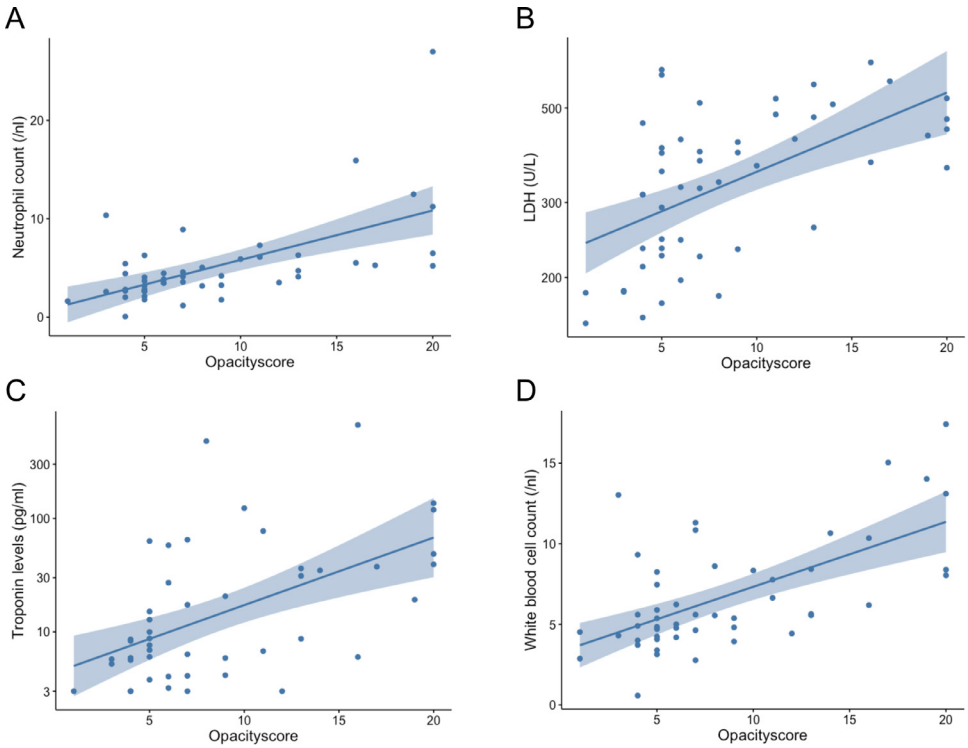

Figure 4. Scatter Plots illustrate correlations between the opacity score and (a) NEU (R = 0.60), (b) LDH (R = 0.60), (c) LEU (R = 0.55) and (d) TNTHS (R = 0.55) from 50 confirmed COVID-19 patients (all $p < 0.001$). (Color version of figure is available online.)

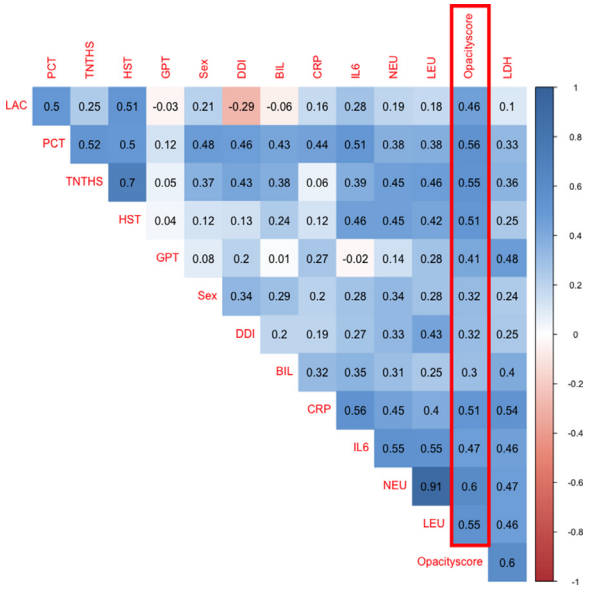

Figure 5. Correlation matrix displays correlations between laboratory data and the opacity score in 50 patients with confirmed COVID-19 infection. Most relevant correlations between laboratory data and the opacity score (all $p < 0.001$) were found between NEU (R = 0.60), LDH (R = 0.6), LEU (R = 0.55), TNTHS (R = 0.55), CRP (R = 0.51). BIL, bilirubin; CRP, C-reactive protein; DDI, d-dimers; () GPT, alanine aminotransferase IL6, interleukin-6; LAC, lactate; LDH, lactate dehydrogenase; (), LEU, white blood cell count; NEU, neutrophil count; TNTHS, troponin; PCT, procalcitonin. (Color version of figure is available online.)

10.15), PCT (2.02 vs. 0.22), IL-6 (375.73 vs. 33.86, $p< 0.01$), CRP (11.64 vs. 5.07, $p< 0.05$) and LDH (419.00 vs. 304.08, $p< 0.05$) (Table 4). Noteworthy, 8 patients ($n = 8$) with mortal disease outcome had a mean number of comorbidities of 1.88 (range, 0-4) and showed a constantly higher opacity score (mean 16.38, range 11-20) compared with survivors ($n = 42$) (mean 7.10, range 1-20) ($p< 0.001$).

DISCUSSION

In this study we evaluated the potential of a state-of-the-art AI-driven software prototype to automatically quantify COVID-19 like lung opacities on chest CT in relationship with course of disease, clinical and laboratory data. The study demonstrated that application of software-based automatic quantification of COVID-19 like opacities is feasible in clinical routine. Furthermore, the statistical analysis revealed moderate to strong correlations between parameters such as the opacity score and length of hospital stay, presence and length of ICU stay, lethal outcome and the severity of laboratory data changes of inflammation parameters, blood count parameters and cardiac markers in 50 laboratory-confirmed COVID-19 patients. ICU ($n = 24$) and non-ICU ($n = 26$) patients differed in several analyses, corroborating the potential of automatic non-invasive CT-scoring to serve as a predictive marker for assessing severity of the clinical course of COVID-19.

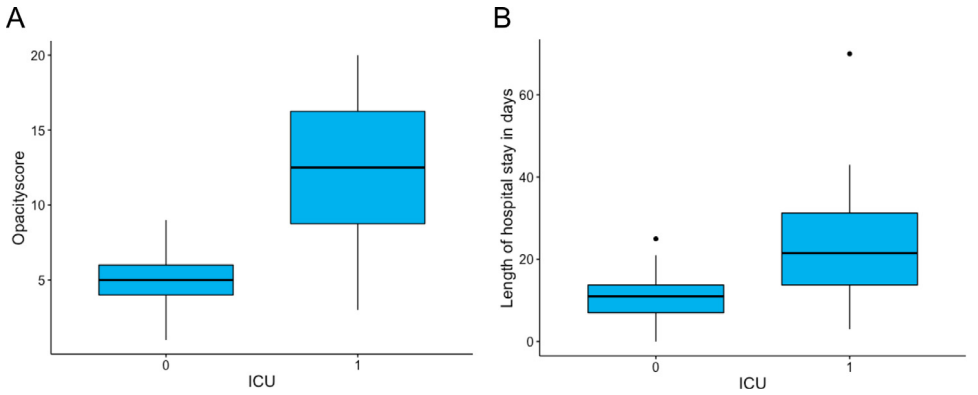

Figure 6. Box-Whisker-Plots illustrate differences regarding (a) the Opacityscore (ranging from 0-20) and (b) Length of hospital stay (in days) between the ICU group and the non-ICU group (all $p < 0.001$). (Color version of figure is available online.)

Clinical workload and organizational effort have continuously increased during the COVID-19 pandemic (17). Non-invasive and rapid screening for potential ICU candidates by applying fully automatically calculated CT-scoring systems may result in a more time-efficient clinical workflow, which may be necessary in the scenario where COVID-19 admission rates rapidly increase (8). In this context, our experience with the prototype's average evaluation time of approximately 2 minutes for a chest 1.0 HR-CT scan emphasizes time-efficient application in clinical routine.

Certain studies have shown that quantitative analysis of COVID-19 like lung findings on chest CT is correlated with clinical and laboratory data and partially might be superior to previous clinical biomarkers in predicting COVID-19 progression to severe illness (8, 18, 19). Nevertheless, manual or semi-automatic quantification of lung lesions on chest CT scans is still prone to inter-observer variance and remains time-consuming (9). In order to reduce bias inherent in semiautomatic analyses we decided to apply a fully automatic software solution to quantify CT scans providing the advantage of a short analysis time and absence of inter-observer variance (14). The prototype applied in our study represents a major step towards transparency and standardization in COVID-19 diagnostics. The software is based on a comprehensible algorithm; the quantification is fully automatic, fast and offers standardized numeric scores in order to quantify the extent of the disease (14). Furthermore, the color-coded illustration of lung lobes enables an easy comprehension and review of the software's analyses. Spatial distribution of opacities can be quickly overviewed by using VRT reconstructions and the high opacity threshold is configurable. Considering that all patients with lethal outcome in this study had a significantly higher opacity score than survivors and the fact that risk models are often based on multiple data which are not directly accessible during an epidemic (20), the use of the software parameters as risk predictors may serve as clinical decision support tool though validity still needs to be elucidated. In addition, the prototype analyses CT scans with a slice thickness of up to 5 mm. Since AI algorithms usually require HR-CT which is frequently associated with consecutive higher radiation and higher costs, the software could be used more widely (e.g. in developing countries) and may contribute to a lower radiation exposure in certain cases (21).

The results of this study are in accordance with previously published results demonstrating that COVID-19 infection contributes to changes in clinical and laboratory parameters - particularly increased CRP, IL-6, LDH and caused lymphopenia - which were significantly associated with disease severity (22-24). DDI and TNTHS showed significant correlations with opacity score in COVID-19 patients, who are known to have an increased risk of acute pulmonary embolism (25) and might have long-term cardiovascular consequence (26). Our findings may indicate that the level of lung manifestation could also be associated with a higher probability of concomitant diseases. Furthermore, this study supports previous findings in terms of association between quantification features and clinical parameters (8, 27). ICU and non-ICU patients significantly differed in opacity score. This is in line with a previously published study describing significant differences in lung opacification percentage among patients with different clinical severity, such as an increased opacity percentage from baseline CT to first follow-up (12). Also, Colombi et al observed that a well-ventilated lung parenchyma of less than 73% on CT on hospital admission was significantly associated with subsequent admission to ICU or death (18). In face of the disadvantage of semi-automatic quantification methods, software-derived parameters might have the potential to predict severity of illness and clinical course and to enable early stratification of critical patients in a fully automatic manner. In addition, the automatic quantification enables an objective evaluation of the COVID-19 like changes independent from the experience of radiologists, which can play a major role in the assessment of the disease, especially for unexperienced observers (28-30).

Gerard et al. used a multi-resolution convolutional neural network to develop a fully automated segmentation algorithm by incorporating CT scans of human with COPD and non-specific lungs of animals with acute lung injury. This lung segmentation was evaluated on CT scans of subjects with COPD, COVID-19, lung cancer and idiopathic pulmonary fibrosis. This study showed that a polymorphic training

TABLE 4. Differences in Clinical and Software Parameters Between Critical and Noncritical Group

Clinical Characteristics	Whole Cohort	Non-Critical cohort	Critical Cohort	*p*-value
Age	57.24 (35-85)	54.58 ± 12.24	60.12 ± 14.67	0.17
Male: female	44:6	20:6	24:0	0.01*
Dyspnea	33 (0.66)	15 (0.58)	18 (0.75)	0.20
Fever	32 (0.64)	21 (0.81)	11(0.46)	0.01*
Cough	28 (0.56)	17 (0.65)	11 (0.46)	0.17
Gastrointestinal symptoms	10 (0.20)	7 (0.27)	3 (0.13)	0.21
Fatigue, head & body aches	12 (0.24)	9 (0.35)	3 (0.13)	0.07
Anosmia & ageusia	8 (0.16)	6 (0.23)	2 (0.08)	0.16
Length of hospital stay	17.22	10.92 ± 5.65	24.04 ± 14.09	<0.001***
No. of comorbidities (median, range)	1.68 (0-5)	1.31 (0-5)	1.79 (0-4)	0.14
Hypertension	17 (0.34)	7 (0.27)	10 (0.42)	0.28
History of cancer	4 (0.08)	1 (0.04)	3 (0.13)	0.27
Diabetes	10 (0.20)	4 (0.15)	6 (0.25)	0.41
Cardiovascular disease	13 (0.26)	4 (0.15)	9 (0.38)	0.08
Lung disease	6 (0.12)	3 (0.12)	3 (0.13)	0.93
Chronic kidney	7 (0.14)	1 (0.04)	6 (0.25)	0.03*
Lethal outcome	8 (0.16)	0	8 (0.33)	<0.01**
Opacity score	8.58 ± 5.34	4.96 ± 1.78	12.50 ± 5.14	<0.001***
Percentage of opacity	28.40 ± 29.70	8.42 ± 8.64	50.06 ± 29.28	<0.001***
Percentage of high opacity	10.46 ± 14.55	2.09 ± 2.39	19.53 ± 16.73	<0.001***
White blood cell count/nl <4.5 >10.0	6.76 ± 3.49 7/50 (0.14) 9/50 (0.18)	5.20 ± 2.49 11/26 (0.42) 1/26 (0.04)	8.45 ± 3.67 2/24 (0.08) 8/24 (0.34)	<0.001***
Lymphocyte count /nl <1.0	0.99 ± 0.53 25/50 (0.5)	1.20 ± 0.60 10/26 (0.38)	0.88 ± 0.40 15/24 (0.63)	0.08
Neutrophil count /nl ≥7.1	5.12 7/48 (0.15)	3.38 ± 12.01 1/25 (0.04)	7.01 ± 5.44 6/23 (0.26)	<0.001***
Neutrophil to lymphocyte ratio ≥3.0	5.32 31/48 (0.65)	3.19 ± 2.32 11/26 (0.42)	7.84 ± 4.49 20/24 (0.83)	<0.001***
C-reactive protein mg/dl ≥10	8.221 ± 7.79 19/50 (0.38)	5.07 ± 5.33 6/26 (0.23)	11.64 ± 8.65 13/24 (0.54)	<0.01**
Interleukin-6 pg/ml ≥7	212.57 41/44 (0.93)	33.86 ± 20.70 18/21 (0.86)	375.73 ± 751.23 23/23 (1.0)	<0.01**
Lactate dehydrogenase U/L ≥245	359.2 ± 135.03 36/50 (0.72)	304.08 ± 124.50 15/26 (0.58)	419.00 ± 121.78 21/24 (0.88)	<0.01**
Troponin pg/ml ≥14	48.32 20/47 (0.43)	10.15 ± 12.02 4/23 (0.17)	84.91 ± 157.47 16/24 (0.67)	<0.001***
D-dimers ng/ml ≥500	2582.6 33/44 (0.75)	964.7 ± 791.86 14/23 (0.61)	4354.52 ± 7841.04 19/21 (0.90)	<0.01**
Lactate mg/dl ≥20	11.26 3/39 (0.07)	9.43 ± 3.58 0/23 (0.0)	13.88 ± 6.75 3/16 (0.19)	0.01*
Procalcitonin ng/ml ≥0.5	1.25 9/40 (0.23)	0.22 ± 0.45 1/17 (0.06)	2.02 ± 3.79 8/23 (0.19)	<0.001***
Bilirubin mg/dl ≥1.4	1.05 ± 2.05 4/50 (0.08)	0.52 ± 0.27 0/26 (0.0)	1.62 ± 2.86 4/24 (0.17)	0.12
Alanine aminotransferase U/L ≥50	61.98 ± 63.39 25/50 (0.5)	53.42 ± 61.94 9/26 (0.)	71.25 ± 64.94 16/24 (0.67)	0.04*
Urea mg/dl ≥50	49.67 12/43 (0.28)	28.95 ± 14.30 2/19 (0.11)	66.08 ± 69.49 10/24 (0.42)	<0.01**
Creatinine mg/dl ≥1.2	1.29 ± 1.14 13/50 (0.26)	0.94 ± 0.24 3/26 (0.12)	1.68 ± 1.55 10/24 (0.25)	0.17
Thrombocyte count /nl <140	206.0 ± 105.78 38/50 (0.76)	208.46 ± 118 6/26 (0.23)	203.33 ± 93.21 6/24 (0.25)	0.85
Quick value % <70	77.54 14/48 (0.29)	79.17 ± 20.45 6/24 (0.25)	75.92 ± 24.05 8/24 (0.33)	0.60

Data are p values, calculated using Wilcoxon-Mann-Whitney-Test.

* Correlation is significant at the 0.05 level.

** Correlation is significant at the 0.01 level.

*** Correlation is significant at the 0.001 level.

enables accurately segmentation of COVID-19 cases and detects common phenotypes without the previous training of data sets with COVID-19 images (31). Harmon et al. used AI for detection of COVID-19 cases and differentiation from other entities in a diverse multi-institutional dataset. The application of a 3D classification and a hybrid 3D model reached an accuracy with high specificity (demonstrating a false negative rate of 10% for other pneumonia entities) and sufficient generalizability (32).

Our study cohort only included COVID-19 laboratory confirmed cases. We are not able to state information about the specificity and sensitivity for the detection of a COVID-19 infection using our software yet. Further, it needs to be investigated, if the technique enables the distinction between different entities of pneumonia.

Since the software works with HU threshold-based detection of opacities it remains unclear whether subtle GGO cannot be detected by the software. Booz et al. showed that MinIP reconstructions could improve the depiction of subtle GGO. Disregarding different slice thicknesses, this method works independently from HU threshold and might lead to an increased sensitivity and therefore to a decline of false-negative cases of non-COVID chest CT scans (33).

Pinkung et al. did an integrative analysis for COVID-19 patient outcome prediction showing that radiomics features describing texture and change of pulmonary opacities in combination with laboratory and demographic data can significantly increase the performance of prediction for a need of ICU admission by an AUC up to 0,884 and sensitivity of 96,1% (34). We found several correlations between image-based severity measures (e.g. opacity score) and likelihood of an ICU admission which might be improved by further inclusion of radiomics features in the future.

This study had several limitations beyond the retrospective study design that need to be addressed. First, correlation does not mean causation. Although COVID-19 might have influenced ICU admission for a certain part, comorbidities could not be deducted. Second, the patient population size was relatively small; studies with larger patient populations are necessary to confirm our results and to further evaluate the potential of the applied software prototype. Third, not all laboratory data were collected on the date of CT performance (±48 h), which may have influenced the results. In addition, the course of laboratory data changes was not included into the correlation analysis in this study. Fourth, CT examinations were carried out at various disease stage due to individual reasons. Fifth, the applied AI-driven software is vendor-specific and currently only applicable on a dedicated research post-processing platform. Sixth, the difference between radiologist and software performance in quantification of COVID-19 like opacities has not been addressed by us in this study.

In this initial study, we demonstrated that fully automatic quantification of COVID-19 like opacities using a software prototype is feasible in clinical routine. The opacity score correlated with clinical and laboratory data in confirmed COVID-19 patients. Furthermore, ICU and non-ICU patients significantly differed regarding the opacity score. Therefore, the automatically computed software parameters offer several advantages compared with semi-automatic quantification methods and might serve as a predictive marker for clinical course of COVID-19 and enable early non-invasive stratification of patients with prognostic and therapeutic relevance. Prospective multicenter studies are required to evaluate software robustness and to further validate translation into clinical practice.

FUNDING

This research did not receive any specific grant from funding agencies in the public, commercial, or non-for-profit sectors.

REFERENCES

1. Zu ZY, Jiang MD, Xu PP, et al. Coronavirus disease 2019 (COVID-19): a perspective from china. Radiology 2020; 296(2):E15–E25.
2. WHO. In: WHO coronavirus disease (COVID-19) dashboard, Geneva: World Health Organization; 2021. Available at https://covid19.who.int Accessed February 13, 2021.
3. Huang C, Wang Y, Li X, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet 2020; 395(10223):497–506.
4. Ai T, Yang Z, Hou H, et al. Correlation of chest CT and RT-PCR testing for coronavirus disease 2019 (COVID-19) in china: a report of 1014 cases. Radiology 2020; 296(2):E32–E40.
5. Prokop M, van Everdingen W, van Rees Vellinga T, et al. CO-RADS—a categorical CT assessment scheme for patients suspected of having COVID-19: definition and evaluation. Radiology 2020; 296(2):E97–E104.
6. Chung M, Bernheim A, Mei X, et al. CT imaging features of 2019 novel coronavirus (2019-nCoV). Radiology 2020; 295(1):202–207.
7. Xiong Y, Sun D, Liu Y, et al. Clinical and high-resolution CT features of the COVID-19 infection: comparison of the initial and follow-up changes. Invest Radiol 2020; 55(6):332–339.
8. Lanza E, Muglia R, Bolengo I, et al. Quantitative chest CT analysis in COVID-19 to predict the need for oxygenation support and intubation. Eur Radiol 2020; 30(12):6770–6778.
9. Shan F, Gao Y, Wang J, et al. Lung infection quantification of COVID-19 in CT. Images Deep Learn 2020: 1–23. arXiv:2003.04655.
10. Bai HX, Wang R, Xiong Z, et al. AI augmentation of radiologist performance in distinguishing COVID-19 from pneumonia of other etiology on chest CT. Radiology 2020; 296:E156–E165.
11. Li L, Qin L, Xu Z, et al. Using artificial intelligence to detect COVID-19 and community-acquired pneumonia based on pulmonary CT: evaluation of the diagnostic accuracy. Radiology 2020; 296(2):E65–E71.
12. Huang L, Han R, Ai T, et al. Serial quantitative chest CT assessment of COVID-19: deep-learning approach. Radiol Cardiothorac Imaging 2020; 2(2):e200075.
13. Wynants L, Van Calster B, Collins GS, et al. Prediction models for diagnosis and prognosis of covid-19 infection: systematic review and critical appraisal. BMJ 2020; 369:m1328.
14. Chaganti S, Grenier P, Balachandran A, et al. Automated quantification of CT patterns associated with COVID-19 from chest CT. Radiol ArtifIntell 2020; 2:e200048.
15. Ghesu FC, Georgescu B, Zheng Y, et al. Multi-scale deep reinforcement learning for real-time 3D-landmark detection in CT scans. IEEE Trans Pattern Anal Mach Intell 2019; 41:176–189.
16. Yang D, Xu D, Zhou S, et al. Automatic liver segmentation using an adversarial image-to-image network. MICCAI2017.
17. Wu X, Hui H, Niu M, et al. Deep learning-based multi-view fusion model for screening 2019 novel coronavirus pneumonia: a multicentre study. Eur J Radiol 2020; 128:109041.
18. Colombi D, Bodini FC, Petrini M, et al. Well-aerated lung on admitting chest CT to predict adverse outcome in COVID-19 pneumonia. Radiology 2020; 296:E86–E96.
19. Liu F, Zhang Q, Huang C, et al. CT quantification of pneumonia lesions in early days predicts progression to severe illness in a cohort of COVID-19 patients. Theranostics 2020; 10:5613–5622.

20. Knight SR, Ho A, Pius R, et al. Risk stratification of patients admitted to hospital with covid-19 using the ISARIC WHO clinical characterisation protocol: development and validation of the 4C mortality score. BMJ 2020; 370:m3339.
21. Li Z, Zhong Z, Li Y, et al. From community-acquired pneumonia to COVID-19: a deep learning-based method for quantitative analysis of COVID-19 on thick-section CT scans. Eur Radiol 2020; 30(12):6828–6837.
22. Zhang ZL, Hou YL, Li DT, et al. Laboratory findings of COVID-19: a systematic review and meta-analysis. Scand J Clin Lab Invest 2020; 80:441–447.
23. Xu YH, Dong JH, An WM, et al. Clinical and computed tomographic imaging features of novel coronavirus pneumonia caused by SARS-CoV-2. J Infect 2020; 80:394–400.
24. Wang D, Hu B, Hu C, et al. Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus-infected pneumonia in Wuhan, China. JAMA 2020; 323(11):1061–1069.
25. Poyiadji N, Cormier P, Patel PY, et al. Acute pulmonary embolism and COVID-19. Radiology 2020; 297(3):E335–E338.
26. Puntmann VO, Carerj ML, Wieters I, et al. Outcomes of cardiovascular magnetic resonance imaging in patients recently recovered from coronavirus disease 2019 (COVID-19). JAMA Cardiol 2020; 5 (11):1265–1273.
27. Zhang K, Liu X, Shen J, et al. Clinically applicable AI system for accurate diagnosis, quantitative measurements, and prognosis of COVID-19 pneumonia using computed tomography. Cell 2020; 182:1360.
28. Wang Y, Chen Y, Wei Y, et al. Quantitative analysis of chest CT imaging findings with the risk of ARDS in COVID-19 patients: a preliminary study. Ann Transl Med 2020; 8:594.
29. Wasilewski PG, Mruk B, Mazur S, Poltorak-Szymczak G, Sklinda K, Walecki J. COVID-19 severity scoring systems in radiological imaging - a review. Pol J Radiol 2020; 85. e361-e8.
30. Das D, Santosh KC, Pal U. Truncated inception net: COVID-19 outbreak screening using chest X-rays. Phys Eng Sci Med 2020; 43:915–925.
31. Gerard SE, Herrmann J, Xin Y, et al. CT image segmentation for inflamed and fibrotic lungs using a multi-resolution convolutional neural network. Sci Rep 2021; 11:1455.
32. Harmon SA, Sanford TH, Xu S, et al. Artificial intelligence for the detection of COVID-19 pneumonia on chest CT using multinational datasets. Nat Commun 2020; 11:4080.
33. Booz C, Vogl TJ, Joseph Schoepf U, et al. Value of minimum intensity projections for chest CT in COVID-19 patients. Eur J Radiol 2021; 135:109478.
34. Chao H, Fang X, Zhang J, et al. Integrative analysis for COVID-19 patient outcome prediction. Med Image Anal 2021; 67:101844.