Korean Journal of Radiology

Prediction of Patient Management in COVID-19 Using Deep Learning-Based Fully Automated Extraction of Cardiothoracic CT Metrics and Laboratory Findings

Thomas Weikert[1], Saikiran Rapaka[2], Sasa Grbic[2], Thomas Re[2], Shikha Chaganti[2], David J. Winkel[1], Constantin Anastasopoulos[1], Tilo Niemann[3], Benedikt J. Wiggli[4], Jens Bremerich[1], Raphael Twerenbold[5], Gregor Sommer[1], Dorin Comaniciu[2], Alexander W. Sauter[1]

Departments of [1]Radiology and [5]Cardiology, University Hospital Basel, University of Basel, Basel, Switzerland; [2]Siemens Healthineers, Princeton, NJ, USA; Departments of [3]Radiology and [4]Infectious Diseases & Infection Control, Kantonsspital Baden, Baden, Switzerland

Objective: To extract pulmonary and cardiovascular metrics from chest CTs of patients with coronavirus disease 2019 (COVID-19) using a fully automated deep learning-based approach and assess their potential to predict patient management.
Materials and Methods: All initial chest CTs of patients who tested positive for severe acute respiratory syndrome coronavirus 2 at our emergency department between March 25 and April 25, 2020, were identified (n = 120). Three patient management groups were defined: group 1 (outpatient), group 2 (general ward), and group 3 (intensive care unit [ICU]). Multiple pulmonary and cardiovascular metrics were extracted from the chest CT images using deep learning. Additionally, six laboratory findings indicating inflammation and cellular damage were considered. Differences in CT metrics, laboratory findings, and demographics between the patient management groups were assessed. The potential of these parameters to predict patients' needs for intensive care (yes/no) was analyzed using logistic regression and receiver operating characteristic curves. Internal and external validity were assessed using 109 independent chest CT scans.
Results: While demographic parameters alone (sex and age) were not sufficient to predict ICU management status, both CT metrics alone (including both pulmonary and cardiovascular metrics; area under the curve [AUC] = 0.88; 95% confidence interval [CI] = 0.79–0.97) and laboratory findings alone (C-reactive protein, lactate dehydrogenase, white blood cell count, and albumin; AUC = 0.86; 95% CI = 0.77–0.94) were good classifiers. Excellent performance was achieved by a combination of demographic parameters, CT metrics, and laboratory findings (AUC = 0.91; 95% CI = 0.85–0.98). Application of a model that combined both pulmonary CT metrics and demographic parameters on a dataset from another hospital indicated its external validity (AUC = 0.77; 95% CI = 0.66–0.88).
Conclusion: Chest CT of patients with COVID-19 contains valuable information that can be accessed using automated image analysis. These metrics are useful for the prediction of patient management.
Keywords: COVID-19; Computed tomography; Deep learning; Artificial intelligence; Patient management

INTRODUCTION

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has caused a global pandemic with over 1.28 million deaths worldwide as of November 12, 2020 [1].

The associated infectious disease, named coronavirus disease 2019 (COVID-19), progresses mildly in most cases [2]. However, severe and critical courses of the disease occur in approximately 20% of all patients [3], mostly demonstrating atypical pneumonia [4]. These patients

require hospitalization or even intensive care unit (ICU) treatment. There are regional differences in utilization of these scarce ressources during a pandemic with temporary shortages. Therefore, criteria for early prediction of patient management, especially whether ICU care is needed or not, are important.

While viral testing remains the only specific method of diagnosis [5], CT plays a role in the workup of suspected pulmonary manifestations of COVID-19 and associated complications. There is growing evidence that radiographic [6] and chest CT [7-15] features are associated with disease severity in COVID-19 based on (semi)-manual assessment and visual scoring of pulmonary parameters. This study intends to build on these approaches and expand them in three aspects: First, by introducing a fully automated and user-independent evaluation method, which is especially relevant in pandemic period with heavy workloads on healthcare providers. Second, this study explicitly focusses on the ultimate patient management status defined by a patient's clinical pathway established with sufficient temporal distance. Third, the inclusion of five cardiovascular metrics covered in all chest CTs has rarely been reported systematically. Notably, preexisting cardiovascular disease is a major risk factor for adverse outcomes in COVID-19 [16]. Laboratory findings were included to assess the value added by the CT metrics.

We hypothesized that CT metrics representing pulmonary and cardiovascular diseases were associated with ultimate patient management in patients with COVID-19 and could help in predicting patient management. It is the goal of this study to extract these CT metrics using a fully automated deep learning-based approach and assess their potential, alone and in combination with laboratory findings and demographic data, for the prediction of patient management.

MATERIALS AND METHODS

This study was approved by the local ethics committee (Ethikkommission Nordwest-und Zentralschweiz; IRB approval number: 2020-00566). It is part of a research project registered on ClinicalTrials.gov on, April 04/29/2020 (Identifier: NCT04366765).

Study Population

All reverse-transcription polymerase chain reaction (RT-PCR) results for SARS-CoV-2 performed at the emergency department (ED) of our institution between March 25 and April 25, 2020, were downloaded from our laboratory data

system (n = 6080 RT-PCR results in 5120 patients). RT-PCR for SARS-CoV-2 was performed using specimens from nasopharyngeal and oropharyngeal swabs. All patients with positive RT-PCR results for COVID-19 were identified (n = 438). In cases with multiple RT-PCRs, the patient was rated positively if a minimum of one of the specimens was positive. For the 438 patients, we searched our RIS/PACS system for the chest CTs performed during the study period, which resulted in 169 chest CTs. At our institution, chest CT is the imaging standard for verifying suspected pulmonary involvement in patients with SARS-CoV-2. For ensuring the independence of observations, all follow-up CTs from a given patient were excluded from the analysis (n = 49). This resulted in 120 CT scans in 120 patients. The time interval between the presentation at the ED and CT acquisition was determined. Figure 1 illustrates the search strategy.

Definition of Patient Management

Information on the ultimate clinical pathway of a patient was retrieved from our hospital information system 12 weeks after the completion of CT data collection (date of determination of ultimate patient management: July 20, 2020). Based on this information, the following three groups were defined: group 1 (outpatient treatment), group 2 (inpatient treatment, general ward), and group 3 (admission to ICU). Each patient was assigned to the highest category individually reached (for instance, a patient that had initially been treated in the general ward and eventually needed ICU care was assigned to group 3).

CT Acquisition Parameters

Chest CT scans were acquired in the supine position using two 128-slice scanners: SOMATOM Definition AS+ (n = 119) and SOMATOM Force (n = 1) (both Siemens Healthineers). Mean tube voltage was 105.0 kVp (standard deviation [SD]: 10.1), mean tube current-time product 81.1 mAs (SD: 19.2), and pitch factor 1.05 in all cases. Most of the scans were performed without a contrast agent (n = 99), whereas 21 CTs were performed with a mean of 71.8 mL (SD: 17.2) of contrast agent (Iopromide, Bayer AG) at an injection rate of 4 mL/s for excluding pulmonary embolism. Images reconstructed in 1-mm slice thickness using soft-tissue reconstruction kernel served as input to the algorithms.

Laboratory Findings

For all patients, the results of six standard laboratory parameters of inflammation and tissue damage were
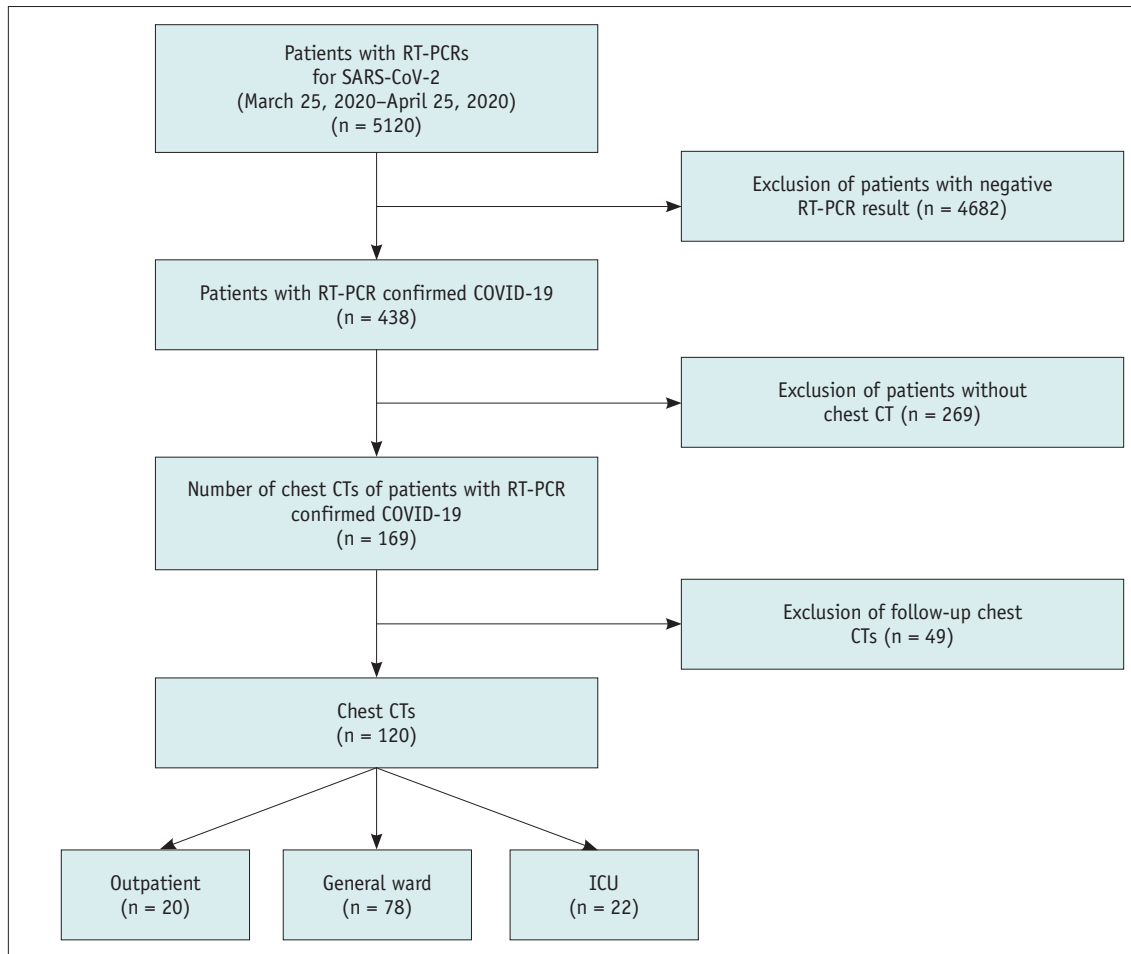
Fig. 1. Search strategy for the main analysis dataset. COVID-19 = coronavirus disease 2019, ICU = intensive care unit, RT-PCR = reverse-transcription polymerase chain reaction, SARS-CoV-2 = severe acute respiratory syndrome coronavirus 2

retrieved from our laboratory system (blood sample type in parentheses): C-reactive protein (CRP; heparin plasma), lactate dehydrogenase (heparin plasma), white blood cell count (EDTA), procalcitonin (heparin plasma), albumin (heparin plasma), and D-dimers (citrate plasma). Laboratory results were obtained on the day of chest CT acquisition.

Technical Details of the Algorithms

Multiple deep convolutional neural networks (DCNNs) were locally deployed on an imaging post-processing platform (Siemens Healthineers, Corporate Technology).

Pulmonary Metrics

The 1-mm series in soft kernel reconstruction served as input to an algorithm prototype based on a deep image-to-image network for lung and lung lobe segmentation and a subsequent DenseUNet for immediate segmentation of opacities. They were trained on chest CTs of n = 9549 (Deep-

Image-to-Image Network) and 901 (DenseUNet) patients, completely independent of the testing dataset used in this study. DenseUNet defined all voxels with ground-glass opacity (GGO) or consolidation as positive/foreground and all other areas of the lung as negative/background. Subsequently, an Hounsfield unit (HU) threshold of -200 was applied to the prediction mask for differentiating GGO from consolidations. Table 1 provides details of all the metrics. Further technical details and high diagnostic performance of the algorithms have been reported previously [17].

Cardiovascular Metric

The non-electrocardiogram-gated 1-mm series served as the only input to a DCNN based on the U-Net architecture for segmentation of the thoracic aorta and the total pericardial volume (TPV). TPV segmentation, which includes the heart and pericardial structures, such as fat and (if present) pericardial effusion, was used to identify

candidate coronary calcification voxels by applying a threshold of > 130 HU. The calcium detection model based on ResNet subsequently predicts the true coronary calcifications. The diameters of the aorta were computed at key anatomical landmarks. The cardiovascular algorithms were trained using 3550 CT scans. Detailed information has been provided elsewhere [18,19]. The quantification of coronary calcifications (QCCs) could only be calculated for series without contrast (n = 99/120). Table 1 lists all the cardiovascular metrics analyzed in this study.

Statistical Analysis

Categorical variables were expressed as counts and percentages. For continuous variables, means with corresponding SDs are provided as measures of variance. For comparing the differences between groups, one-way analyses of variance for normally distributed continuous variables, the Kruskal-Wallis H tests for non-normally distributed continuous variables, and the chi-square tests for categorical variables were performed. The statistical analysis comprised the following three steps:

Step 1: A Series of Univariable Analyses with Appropriate post hoc Tests to assess the association of CT metrics, laboratory findings, and patient characteristics with patient management. Studies with contrast were excluded during the analysis of QCC, as this measure was only calculated on a non-contrast series.

Step 2: A Series of Multivariable Binary Logistic Regressions to assess the potential of CT metrics, laboratory findings, and patient demographics as well as their combinations, to classify patients who needed ICU care from those who did not. The ICU status of a patient (0 = no ICU; 1 = ICU) served as the dependent variable. CT metrics, laboratory findings, and patient demographics (age and sex) served as independent variables. Inclusion criteria for CT metrics and laboratory findings were p values ≤ 0.05, in the subgroup comparisons between groups 1 and 3 (outpatient vs. ICU) or group 2 vs. 3 (general ward vs. ICU) in Step 1 of the analysis. Furthermore, the parameters had to be available for all patients. This resulted in the following five models:

- D: Patients demographics only
- L: Laboratory findings only
- PC: CT metrics only
- PD: Pulmonary CT metrics and demographics*
- PCLD: All parameters (CT metrics, laboratory findings, demographics)

*This model was created for external validation (details below).

For all approaches, area under the curve (AUC) with 95% confidence intervals (CIs) were calculated using prediction

Table 1. Features Analyzed by Algorithms with Units and Definitions

Metric Name	Unit	Definition
Pulmonary metrics		
Lung volume	mL	Total volume of the lung
PO	%	Percentage of volume of the lung affected by opacities, equivalent to GGO and consolidation
PHO	%	Percentage of volume of the lung affected by high HU opacities (equal to or above -200 HU), equivalent to consolidations
%LowHU	%	Percentage of volume of the lung with low attenuation, defined as < -950 HU
LL#	N	Number of lung lesions detected (GGO or consolidations)
LSS	N	Sum of severity scores for each of the five lung lobes*
LHOS	N	Sum of severity scores for each of the five lung lobes, consolidation only*
Cardiovascular metrics		
TPV	mL	Total pericardial volume. This includes the heart as well as pericardial structures like fat and/or effusion
QCC	mm³	Total volume of coronary calcifications. Only available for CT without contrast
D_AAsc	mm	Diameter of ascending aorta at the level of the right pulmonary artery
D_Arch	mm	Diameter of aorta, midway through aortic arch region
D_ADsc	mm	Diameter of descending aorta, midway between the branching of the subclavian artery and celiac artery

*Severity scores calculated as follows. 0: lobe not affected by GGO/consolidation, 1: 1–25% of lobe volume affected, 2: 25–50%, 3: 50–75%, 4: 75–100%. GGO = ground-glass opacity, HU = Hounsfield unit, QCC = quantification of coronary calcification

probabilities obtained from the binary logistic regression analyses. Furthermore, we analyzed the differences in the AUCs between the models according to the method proposed by DeLong et al. [20].

Step 3: Internal and External Validation to assess generalizability. For internal validation, we processed all chest CTs of patients with positive RT-PCR for SARS-CoV-2 acquired at our institution during a later period (April 26, 2020–May 20, 2020; information on ultimate patient management retrieved on August 12, 2020) with all algorithms. For testing external validity, we used data from another hospital (Supplementary Materials, Supplementary Table 1). ICU status was predicted using regression equations obtained from Step 2 of the analysis.

Statistical analyses were performed with IBM SPSS Statistics for Windows, Version 22.0 (IBM Corp.), using default settings. The p values ≤ 0.05 indicated statistical significance.

RESULTS

Patient Characteristics

The main analysis dataset of this study included 120 patients with a mean age of 60.8 years (SD: 17.5; range: 18–92 years; 47 [39.2%] females). Table 2 summarizes the patient characteristics of the three patient management groups.

Automated CT Metrics Analysis

All the datasets were successfully processed using the algorithm. The mean time interval between presentation at ED and CT acquisition was 0.98 days (SD: 2.32 days).

Step 1: Association of Metrics with Patient Management

Table 3 summarizes the results of the univariable analyses of CT metrics and laboratory findings.

Pulmonary CT Metrics

PO, PHO, LSS, and LHOS steadily increased continuously from group 1 to group 3, while lung volume and %LowHU decreased from group 1 to group 3. All these differences were statistically significant. Post hoc testing revealed that differences in PO, PHO, LSS, and LHOS were statistically significant at p values < 0.01 between all three subgroups. Regarding lung volume and %LowHU, group comparisons 1 vs. 3 and 2 vs. 3 differed statistically significantly. Figure 2 displays an image example for each group.

Cardiovascular CT Metrics

TPV and D_AAsc differed significantly among the three groups. Post hoc analysis revealed that differences in both TPV and D_AAsc were statistically significant only for the comparison of groups 1 and 3 (TPV: p = 0.041; D_AAsc: p = 0.033). QCC, D_Arch, and D_ADsc did not differ significantly. Figure 3 illustrates the outputs of the cardiovascular algorithms.

Laboratory Findings

Laboratory analysis for D-dimers and procalcitonin was not performed in some cases (11/120 and 36/120, respectively). All laboratory parameters differed significantly among the three groups. CRP levels increased steadily from groups 1 to 3, while albumin decreased. Subgroup comparisons were statistically significant for all group comparisons (CRP, albumin), group comparisons 1 vs. 3, and 2 vs. 3 (lactate dehydrogenase and procalcitonin), group comparison 2 vs. 3 only (white blood cell count), and group comparison 1 vs. 3 only (D-dimer).

Step 2: Prediction of ICU Status

Table 4 specifies metrics and parameters included in the five multivariable models for classification of ICU status (yes/no) according to the criteria mentioned in the methods section. The best performing model was the PCLD model combining CT-derived, laboratory, and demographic parameters (AUC = 0.91). Demographic parameters alone could not distinguish ICU patients from non-ICU patients (AUC = 0.55). CT-derived metrics (including both pulmonary and cardiovascular metrics) alone, laboratory metrics alone,

Table 2. Patient Characteristics in Patient Management Groups with Intergroup Comparison

	Group 1 (Outpatient)	Group 2 (General Ward)	Group 3 (ICU)	*P* (Intergroup Comparison)
Number of patients	20	78	22	
Mean age ± SD, year	58.2 ± 19.3	62.1 ± 17.4	63.1 ± 14.6	0.100
Sex, % female	35.0 (7/20)	41.0 (32/78)	36.4 (8/22)	0.847

ICU = intensive care unit, SD = standard deviation

Table 3. Results of Univariable Analyses of CT Metrics and Laboratory Findings

Parameter	Unit	Group 1	Group 2	Group 3	Statistic	P
Pulmonary metrics						
Lung volume	mL	4917.4 (1286.2)	4336.5 (1205.5)	3611.3 (1059.5)	ANOVA	0.002*
P0	%	3.7 (6.1)	15.6 (15.9)	39.4 (24.1)	Kruskal-Wallis	< 0.001*
PH0	%	0.6 (1.2)	3.3 (4.5)	10.1 (8.6)	Kruskal-Wallis	< 0.001*
%LowHU	%	7.8 (6.8)	5.4 (4.3)	2.7 (2.4)	Kruskal-Wallis	< 0.001*
LL#	N	17.3 (10.9)	28.7 (17.3)	21.6 (10.7)	ANOVA	0.061
LSS	N	2.3 (2.2)	5.7 (3.5)	10.8 (4.8)	Kruskal-Wallis	< 0.001*
LHOS	N	0.9 (1.3)	2.8 (2.0)	5.1 (2.2)	Kruskal-Wallis	< 0.001*
Cardiovascular metrics						
TPV	mL	762.1 (194.5)	881.3 (236.6)	908.8 (166.1)	Kruskal-Wallis	0.043*
QCC	mm^2	321.4 (700.3)	335.8 (633.6)	461.9 (915.6)	ANOVA	0.552
D_AAsc	mm	32.9 (3.9)	35.5 (4.4)	36.6 (4.1)	Kruskal-Wallis	0.038*
D_Arch	mm	28.8 (3.5)	29.6 (2.9)	30.4 (2.7)	Kruskal-Wallis	0.184
D_ADsc	mm	25.0 (4.3)	27.2 (2.6)	26.0 (4.3)	Kruskal-Wallis	0.215
Laboratory findings						
CRP	mg/L	32.5 (59.9)	62.0 (73.5)	104.5 (81.9)	Kruskal-Wallis	< 0.001*
Lactate de-hydrogenase	U/L	232.7 (84.6)	283.0 (159.7)	406.8 (144.0)	Kruskal-Wallis	< 0.001*
White blood cell count	10^9/L	6.9 (2.9)	6.9 (3.4)	9.4 (4.3)	Kruskal-Wallis	0.035*
Procalcitonin	μ/L	0.099 (0.141)	0.205 (0.531)	0.399 (0.427)	Kruskal-Wallis	< 0.001*
Albumin	g/L	35.8 (4.5)	30.8 (6.3)	25.32 (6.4)	Kruskal-Wallis	< 0.001*
D-Dimer	μ/mL	0.472 (0.302)	1.113 (1.438)	1.397 (1.401)	Kruskal-Wallis	0.016*

Data are mean values and standard deviations (in brackets). *$p \leq 0.05$. CRP = C-reactive protein, QCC = quantification of coronary calcification, TPV = total pericardial volume

and pulmonary CT metrics combined with demographic parameters were all good classifiers with AUCs ≥ 0.84. Table 5 provides detailed information. The AUC of the D model differed significantly from that of all other models ($p < 0.001$). The difference in the AUCs of models with CT-derived parameters alone vs. laboratory parameters alone was not statistically significant ($p = 0.462$). Figure 4 displays the receiver operating characteristic curves of the PCLD, PC, and L models. As D-dimers and procalcitonin were not available in all cases, these two parameters were excluded from the analysis.

Step 3: Internal and External Validation

The internal validation comprising 16 new cases of patients with positive RT-PCR results for SARS-CoV-2 resulted in a sensitivity of 80.0% (4 of 5 patients admitted to ICU correctly classified) and a specificity of 81.8% (9 of 11 patients not admitted to ICU correctly classified) for the PCLD model. Table 5 presents the results. In general, the performance measures on the internal validation dataset were slightly worse than those on the main analysis dataset but still acceptable. The mean age of the internal validation dataset was 63.0 years (SD: 16.0) and not statistically

significantly different from the dataset used for the main analysis ($p = 0.635$). We also found evidence for external validity using the PD model (Supplementary Materials for detailed information). Table 6 demonstrates that demographic information did not differ significantly among the three datasets.

DISCUSSION

This study demonstrated that it is feasible to automatically extract pulmonary and cardiovascular metrics from chest CT scans of patients with RT-PCR-confirmed COVID-19 that are useful for the prediction of patient management. Multiple CT metrics continuously and significantly increased or decreased with intensified patient management. The same was true for laboratory parameters reflecting inflammation and cell damage. The best prediction regarding ICU status was achieved by combining CT metrics, laboratory findings, and demographic information, while the latter alone could not differentiate the two classes. The CT metrics and laboratory findings were good classifiers on their own. Although internal and external validation demonstrated marginally inferior
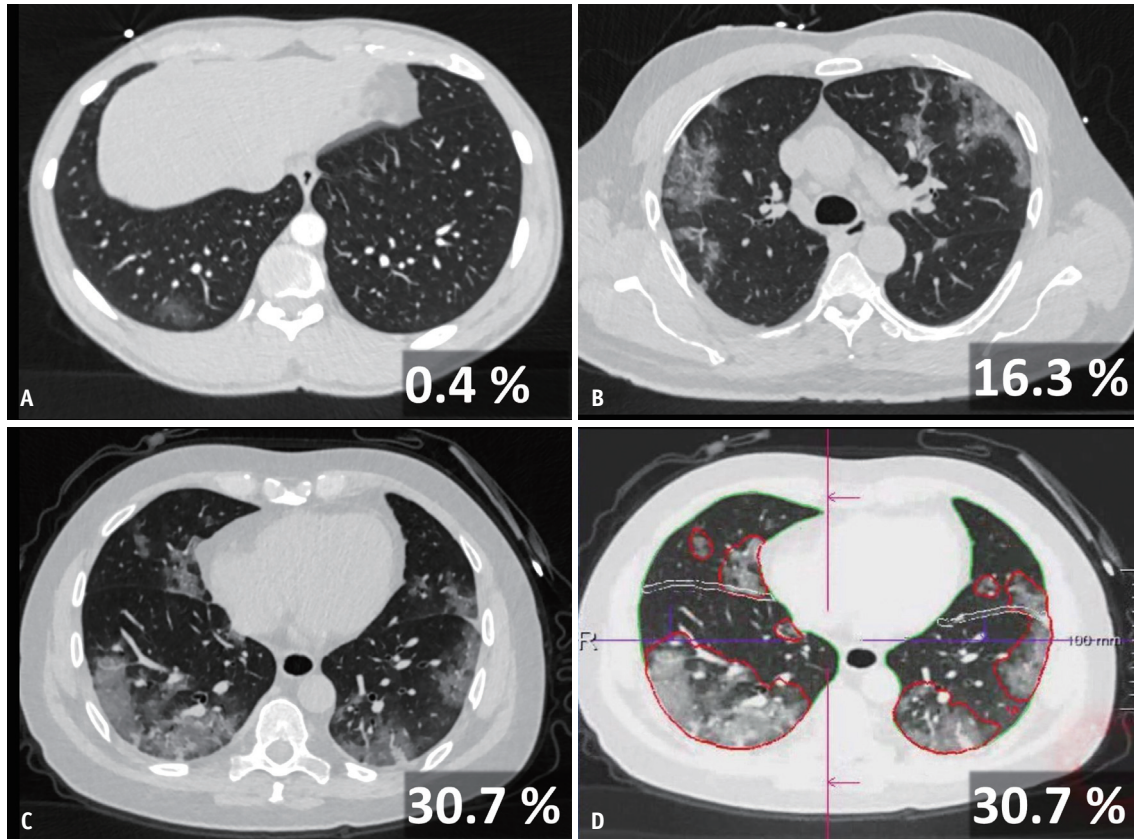
Fig. 2. Exemplary axial chest CT images for groups 1 (A), 2 (B), and 3 (C) with PO in percent in the lower right corner. (D) shows the output of pulmonary opacity segmentation algorithm for case (C), with fine delineation of areas with pulmonary opacities in red color.
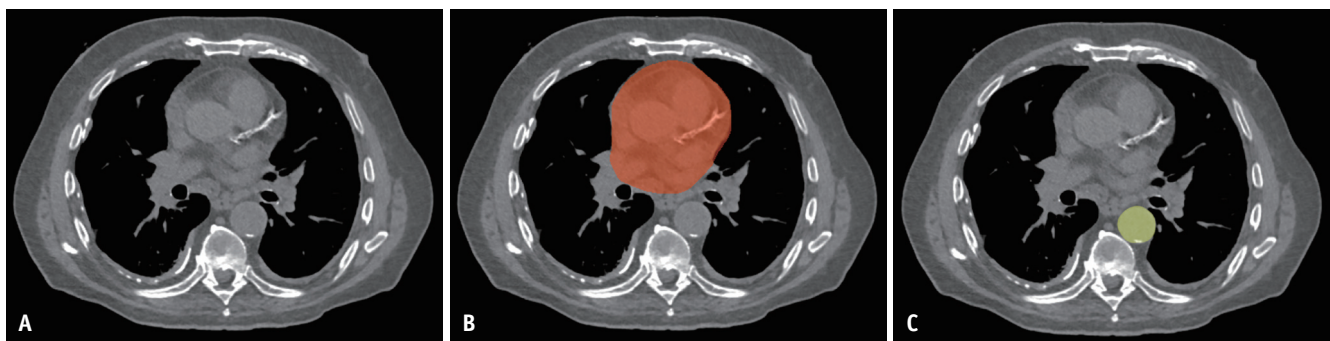

Fig. 3. Exemplary axial chest CT images visualizing the output of cardiovascular algorithms: original input image (A), segmentation of the total pericardial volume as a red overlay (B), and segmentation of the descending aorta as a yellow overlay (C) as the basis for diameter calculations.

Table 4. Metrics Included in Five Multivariable Models for Prediction of Patient Management Regarding Need of ICU-Care (Yes/No)

Model	Metrics and Parameters Included
D	Age, sex
L	CRP, lactate dehydrogenase, white blood cell count, albumin
PC	Lung volume, PO, PHO, %LowHU, LSS, LHOS, TPV, D_AAsc
PD	Lung volume, PO, PHO, age, sex
PCLD	Lung volume, PO, PHO, %LowHU, LSS, LHOS, TPV, D_AAsc, CRP, lactate dehydrogenase, white blood cell count, albumin, age, sex

Name of models indicates which metric/parameter groups were included as independent variables. C = cardiovascular metrics, CRP = C-reactive protein, D = demographic paramters, ICU = intensive care unit, L = laboratory findings, P = pulmonary CT metrics

performance, it was still good.

Our results regarding the relevance of pulmonary CT metrics in COVID-19 and their association with patient management are in line with previous studies and expected, as they reflect pathologic changes, concretely inflammatory GGO, and consolidations. Li et al. [8] reported an increasing extent of inflammatory pulmonary lesions from light to common to severe/critical clinical manifestations. Sun et

al. [7] and Tan et al. [21] confirmed that quantitative CT parameters strongly correlate with laboratory inflammation markers. Lyu et al. [22] showed that the number of lung segments and lobes affected by consolidations increased with case severity, which is in line with the increase in LSS and LHOS with higher admission status [22]. Similarly, Liu et al. [10] reported an association between a higher lung severity score and extended hospitalization time. A

Table 5. Performance Measures of Five Multivariable Models for Prediction of Patient Management Regarding Need of ICU-Care (Yes/No)

Model	Statistic	Main Analysis Dataset	Internal Validation Dataset	External Validation Dataset
D	Sensitivity (%)	63.6 (14/22)	40 (2/5)	NA
	Specificity (%)	50.0 (49/98)	36.4 (4/11)	NA
	AUC (95% CI)	0.55 (0.42-0.67)	0.42 (0.15–0.79)	NA
L	Sensitivity (%)	90.9 (20/22)	80.0 (4/5)	NA
	Specificity (%)	73.5 (72/98)	45.5 (5/11)	NA
	AUC (95% CI)	0.86 (0.77–0.94)	0.47 (0.15–0.79)	NA
PC	Sensitivity (%)	90.9 (20/22)	80.0 (4/5)	NA
	Specificity (%)	78.6 (77/98)	72.7 (8/11)	NA
	AUC (95% CI)	0.88 (0.79–0.97)	0.75 (0.47–1.00)	NA
PD	Sensitivity (%)	81.8 (18/22)	60.0 (3/5)	74.2 (23/31)
	Specificity (%)	70.4 (69/98)	72.7 (8/11)	75.8 (47/62)
	AUC (95% CI)	0.84 (0.75–0.94)	0.71 (0.39–1.00)	0.77 (0.66–0.88)
PCLD	Sensitivity (%)	95.5 (21/22)	80.0 (4/5)	NA
	Specificity (%)	75.5 (74/98)	81.8 (9/11)	NA
	AUC (95% CI)	0.91 (0.85–0.98)	0.75 (0.48–1.00)	NA

Sensitivity and specificity at optimal cutoff according to Youden with corresponding number of patients in brackets. AUC = area under the curve, CI = confidence interval, ICU = intensive care unit, NA = not available
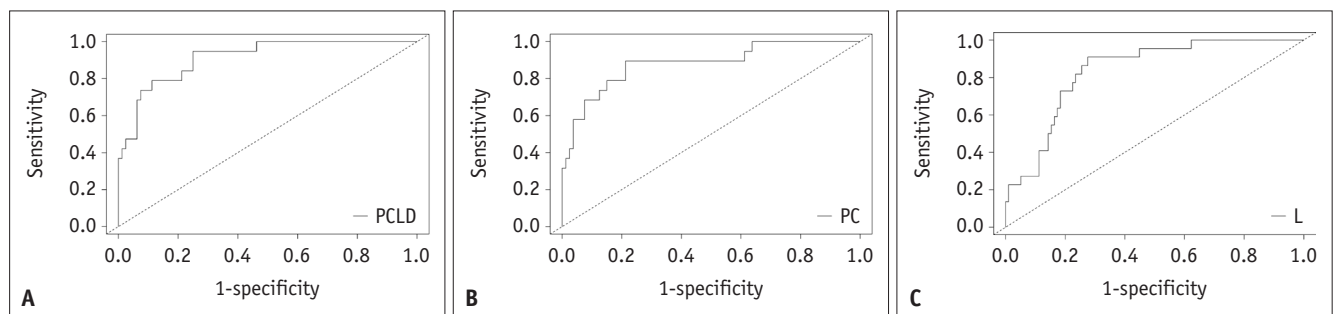

Fig. 4. Receiver operating characteristic curves for the PCLD model (A), PC model (B), and L model (C) to assess discriminatory power regarding ICU-care on main analysis dataset (whether a patient finally required ICU care or not). A combination CT metrics, laboratory findings, and patient demographics was the best classifier (**A;** area under the curve = 0.91; 95% confidence interval = 0.85–0.98). ICU = intensive care unit

Table 6. Comparison of Patient Characteristics in Main Analysis Dataset vs. Internal Validation Dataset vs. External Validation Dataset

	Main Analysis Dataset	Internal Validation	External Validation	P (Intergroup Comparison)
Number of patients	120	16	93	
Mean age ± SD, year	60.8 ± 17.5	63.0 ± 16.5	64.8 ± 16.1	0.678
Sex, % female	39.2 (47/120)	41.0 (4/16)	37.6 (35/93)	0.546

Differences in age and sex were not statistically significantly different between groups, as indicated by the p values.

significant number of additional studies have successfully applied lung volume assessment with or without a combination of clinical and laboratory tests for predicting disease severity, treatment intensity, outcome, and mortality [7,10,15,23-26]. Notably, the analyses in these studies required substantial manual interaction and visual assessment.

However, in a pandemic with limited human resources, fully automated approaches are preferred. In this respect, Huang et al. [27] applied CT-derived opacification measures using deep learning to stratify four clinical subtypes according to their baseline clinical, laboratory, and CT findings. They provided further evidence of CT as an important tool for risk stratification in patients with COVID-19 and reported percentages of lung areas with opacities ranging from 0% (mild disease) to 49.6% (critical disease), which is in line with the results of the analysis at hand. However, radiological findings used to predict the outcomes were at the same time part of the outcome definition criteria of this study [28].

As previously shown, preexisting cardiovascular disease is a risk factor for adverse outcomes in COVID-19 [24] and, COVID-19 simultaneously affects the cardiovascular system [29]. However, the aforementioned approaches did not include quantitative measurements of cardiovascular CT metrics. This study included cardiovascular metrics, such as TPV, as an estimate of heart size. Indeed, a higher TPV was associated with a higher risk of intensified patient management. As age and sex did not differ significantly between groups, differences were caused probably by increased heart size or increased amount of pericardial fat.

The AUCs of the models considering CT-derived metrics only vs. laboratory paramters only were both high and did not differ statistically significantly. This is probably due to the fact that both reflect inflammation of the lungs and are highly correlated. Internal validation indicated good internal generalizability, as did the external data for the PD model.

This study had several limitations. First, the internal validation dataset was small, resulting in wide CIs; therefore, the results should be interpreted cautiously. However, high standardization of chest CT and the fact that other studies on the topic reported similar effect sizes provide confidence that the results are generalizable. Second, while the main investigator site had access to the algorithms with all pulmonary and cardiovascular metrics, the remote site had access to pulmonary metrics only.

Third, this study included only patients with COVID-19 who underwent a CT scan, the diagnostic standard for patients with suspected pulmonary manifestations of COVID-19 at our center. Therefore, the presented approach might be less relevant in medical centers that rarely perform chest CT. Fourth, other features, such as the initial severity of symptoms, might be useful to classify patient management. Besides the focus on automatically retrieved CT metrics, this study also considered demographic parameters and laboratory findings.

To conclude, this study provides evidence that chest CT of patients with COVID-19 contains valuable information for the prediction of ultimate patient management. Furthermore, this information is accessible using a deep learning-based fully automated image analysis workflow, which is especially helpful during the COVID-19 pandemic.

Supplementary Materials

The Data Supplement is available with this article at https://doi.org/10.3348/kjr.2020.0994.

Author Contributions
Conceptualization: Thomas Weikert, Alexander W. Sauter, Jens Bremerich, Dorin Comaniciu. Data curation: Thomas Weikert, Sasa Grbic, Shikha Chaganti, Raphael Twerenbold. Formal analysis: Thomas Weikert, Thomas Re, Tilo Niemann, Benedikt J. Wiggli, Gregor Sommer, Alexander W. Sauter. Investigation: Thomas Weikert, Saikiran Rapaka, Constantin Anastasopoulos, Alexander W. Sauter. Methodology: Thomas Weikert, Saikiran Rapaka, David J. Winkel, Constantin Anastasopoulos, Alexander W. Sauter. Project administration: Thomas Weikert, Alexander W. Sauter.

Resources: Sasa Grbic, Thomas Re, Tilo Niemann, Benedikt J. Wiggli, Jens Bremerich, Raphael Twerenbold, Dorin Comaniciu, Alexander W. Sauter. Software: Shikha Shaganti, Gregor Sommer, Dorin Comaniciu. Supervision: Dorin Comaniciu, Alexander W. Sauter. Validation: Thomas Weikert, Tilo Niemann, Benedikt J. Wiggli, Gregor Sommer, Alexander W. Sauter. Visualization: Thomas Weikert, Shikha Chaganti, Thomas Re, Alexander W. Sauter. Writing—original draft: Thomas Weikert, Alexander W. Sauter. Writing—review & editing: Thomas Weikert, Sasa Grbic, Shikha Chaganti, David J. Winkel, Constantin Anastasopoulos, Raphael Twerenbold, Gregor Sommer, Jens Bremerich, Alexander W. Sauter.

ORCID iDs

Thomas Weikert
    https://orcid.org/0000-0001-9274-053X
Saikiran Rapaka
    https://orcid.org/0000-0001-6527-0974
Sasa Grbic
    https://orcid.org/0000-0002-7564-9811
Thomas Re
    https://orcid.org/0000-0003-2285-1562
Shikha Chaganti
    https://orcid.org/0000-0002-7029-2665
David J. Winkel
    https://orcid.org/0000-0001-7051-8022
Constantin Anastasopoulos
    https://orcid.org/0000-0001-9818-2460
Tilo Niemann
    https://orcid.org/0000-0002-5384-2522
Benedikt J. Wiggli
    https://orcid.org/0000-0002-6197-3530
Jens Bremerich
    https://orcid.org/0000-0002-1002-8483
Raphael Twerenbold
    https://orcid.org/0000-0003-3814-6542
Gregor Sommer
    https://orcid.org/0000-0002-8952-0808
Dorin Comaniciu
    https://orcid.org/0000-0002-5238-8647
Alexander W. Sauter
    https://orcid.org/0000-0002-6707-2258

REFERENCES

1. Johns Hopkins University. COVID-19 Map - Johns Hopkins Coronavirus Resource Center. Coronavirus.jhu.edu Web site. https://coronavirus.jhu.edu/map.html. Accessed May 26, 2020
2. Epidemiology Working Group for NCIP Epidemic Response, Chinese Center for Disease Control and Prevention. The epidemiological characteristics of an outbreak of 2019 novel coronavirus diseases (COVID-19) in China. *Zhonghua Liu Xing Bing Xue Za Zhi* 2020;41:145-151
3. Verity R, Okell LC, Dorigatti I, Winskill P, Whittaker C, Imai N, et al. Estimates of the severity of coronavirus disease 2019: a model-based analysis. *Lancet Infect Dis* 2020;20:669-677
4. Mehta P, McAuley DF, Brown M, Sanchez E, Tattersall RS, Manson JJ; HLH Across Speciality Collaboration, UK. COVID-19: consider cytokine storm syndromes and immunosuppression. *Lancet* 2020;395:1033-1034
5. American College of Radiology. ACR recommendations for the use of chest radiography and computed tomography (CT) for suspected COVID-19 infection. Acr.org Web site. https://www.acr.org/Advocacy-and-Economics/ACR-Position-Statements/Recommendations-for-Chest-Radiography-and-CT-for-Suspected-COVID19-Infection. Published 2020. Accessed August 6, 2020
6. Hwang EJ, Kim H, Yoon SH, Goo JM, Park CM. Implementation of a deep learning-based computer-aided detection system for the interpretation of chest radiographs in patients suspected for COVID-19. *Korean J Radiol* 2020;21:1150-1160
7. Sun D, Li X, Guo D, Wu L, Chen T, Fang Z, et al. CT quantitative analysis and its relationship with clinical features for assessing the severity of patients with COVID-19. *Korean J Radiol* 2020;21:859-868
8. Li K, Fang Y, Li W, Pan C, Qin P, Zhong Y, et al. CT image visual quantitative evaluation and clinical classification of coronavirus disease (COVID-19). *Eur Radiol* 2020;30:4407-4416
9. Zhang R, Ouyang H, Fu L, Wang S, Han J, Huang K, et al. CT features of SARS-CoV-2 pneumonia according to clinical presentation: a retrospective analysis of 120 consecutive patients from Wuhan city. *Eur Radiol* 2020;30:4417-4426
10. Liu Z, Jin C, Wu CC, Liang T, Zhao H, Wang Y, et al. Association between initial chest CT or clinical features and clinical course in patients with coronavirus disease 2019 pneumonia. *Korean J Radiol* 2020;21:736-745
11. 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;27:910-921
12. Li M, Lei P, Zeng B, Li Z, Yu P, Fan B, et al. Coronavirus disease (COVID-19): spectrum of CT findings and temporal progression of the disease. *Acad Radiol* 2020;27:603-608
13. Zheng Y, Xiao A, Yu X, Zhao Y, Lu Y, Li X, et al. Development and validation of a prognostic nomogram based on clinical and CT features for adverse outcome prediction in patients with COVID-19. *Korean J Radiol* 2020;21:1007-1017
14. Yin X, Min X, Nan Y, Feng Z, Li B, Cai W, et al. Assessment of the severity of coronavirus disease: quantitative computed

tomography parameters versus semiquantitative visual score. *Korean J Radiol* 2020;21:998-1006

15. Park B, Park J, Lim JK, Shin KM, Lee J, Seo H, et al. Prognostic implication of volumetric quantitative ct analysis in patients with COVID-19: a multicenter study in Daegu, Korea. *Korean J Radiol* 2020;21:1256-1264

16. Driggin E, Madhavan MV, Bikdeli B, Chuich T, Laracy J, Biondi-Zoccai G, et al. Cardiovascular considerations for patients, health care workers, and health systems during the COVID-19 pandemic. *J Am Coll Cardiol* 2020;75:2352-2371

17. Chaganti S, Grenier P, Balachandran A, Chabin G, Cohen S, Flohr T, et al. Automated quantification of CT patterns associated with COVID-19 from chest CT. *Radiology: Artificial Intelligence* 2020;2:e200048

18. Ali A, Balachandran A, Vishwanath RS, Barthur A, Wichmann JL, Cimen S, et al. Evaluation of a deep learning based aortic diameter quantification system against multi-reader consensus measurement. Proceedings 2020 European Congress of Radiology (ECR); 2020 Jul 15-19; Vienna, Austria: European Society of Radiology

19. Martin SS, van Assen M, Rapaka S, Hudson HT Jr, Fischer AM, Varga-Szemes A, et al. Evaluation of a deep learning-based automated CT coronary artery calcium scoring algorithm. *JACC Cardiovasc Imaging* 2020;13:524-526

20. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 1988;44:837-845

21. Tan C, Huang Y, Shi F, Tan K, Ma Q, Chen Y, et al. C-reactive protein correlates with computed tomographic findings and predicts severe COVID-19 early. *J Med Virol* 2020;92:856-862

22. Lyu P, Liu X, Zhang R, Shi L, Gao J. The performance of chest CT in evaluating the clinical severity of COVID-19 pneumonia: identifying critical cases based on CT characteristics. *Invest Radiol* 2020;55:412-421

23. Francone M, Iafrate F, Masci GM, Coco S, Cilia F, Manganaro L, et al. Chest CT score in COVID-19 patients: correlation with disease severity and short-term prognosis. *Eur Radiol* 2020;30:6808-6817

24. Li B, Yang J, Zhao F, Zhi L, Wang X, Liu L, et al. Prevalence and impact of cardiovascular metabolic diseases on COVID-19 in China. *Clin Res Cardiol* 2020;109:531-538

25. Ruch Y, Kaeuffer C, Ohana M, Labani A, Fabacher T, Bilbault P, et al. CT lung lesions as predictors of early death or ICU admission in COVID-19 patients. *Clin Microbiol Infect* 2020;26:1417.e5-1417.e8

26. Kim YC, Chung Y, Choe YH. Automatic localization of anatomical landmarks in cardiac MR perfusion using random forests. *Biomed Signal Proces* 2017;38:370-378

27. Huang L, Han R, Ai T, Yu P, Kang H, Tao Q, et al. Serial quantitative chest CT assessment of COVID-19: a deep learning approach. *Radiology: Cardiothoracic Imaging* 2020;2:e200075

28. National Health Commission & National Administration of Traditional Chinese Medicine. Diagnosis and treatment protocol for novel coronavirus pneumonia. *Chin Med J (Engl)* 2020;133:1087-1095

29. Zheng YY, Ma YT, Zhang JY, Xie X. COVID-19 and the cardiovascular system. *Nat Rev Cardiol* 2020;17:259-260