

Original Article

Designing the CORI score for COVID-19 diagnosis in parallel with deep learning-based imaging models

Telly Kamelia^{1,2*}, Benny Zulkarnaen^{3,4}, Wita Septiyanti^{3,4}, Rahmi Afifi^{3,4}, Adila Krisnadhi⁵, Cleopas M. Rumende^{1,2}, Ari Wibisono⁵, Gladhi Guarddin⁵, Dina Chahyati⁵, Reyhan E. Yunus^{3,4}, Dhita P. Pratama⁵, Irda N. Rahmawati^{6,7}, Dewi Nareswari^{6,7}, Maharani Falerisya^{6,7}, Raissa Salsabila^{6,7}, Bagus DI. Baruna⁸, Anggraini Iriani⁹, Finny Nandipinto¹⁰, Ceva Wicaksono^{1,2} and Ivan R. Sini¹¹

¹Division of Respiriology and Critical Care, Department of Internal Medicine, Faculty of Medicine, Universitas Indonesia, Jakarta, Indonesia; ²Division of Respiriology and Critical Care, Department of Internal Medicine, Cipto Mangunkusumo National General Hospital, Jakarta, Indonesia; ³Department of Radiology, Faculty of Medicine, Universitas Indonesia, Jakarta, Indonesia; ⁴Department of Radiology, Cipto Mangunkusumo National General Hospital, Jakarta, Indonesia; ⁵Faculty of Computer Science, Universitas Indonesia, Jakarta, Indonesia; ⁶Department of Internal Medicine, Faculty of Medicine, Universitas Indonesia, Jakarta, Indonesia; ⁷Department of Internal Medicine, Cipto Mangunkusumo National General Hospital, Jakarta, Indonesia; ⁸Department of Radiology, Bunda Jakarta General Hospital, Jakarta, Indonesia; ⁹Department of Clinical Pathology, Bunda Jakarta General Hospital, Jakarta, Indonesia; ¹⁰Department of Radiology, Bunda Margonda General Hospital, Jakarta, Indonesia; ¹¹IRSI Research and Training Centre, Jakarta, Indonesia

*Corresponding author: tellykamelia99@gmail.com

Abstract

The coronavirus disease 2019 (COVID-19) pandemic has triggered a global health crisis and placed unprecedented strain on healthcare systems, particularly in resource-limited settings where access to RT-PCR testing is often restricted. Alternative diagnostic strategies are therefore critical. Chest X-rays, when integrated with artificial intelligence (AI), offers a promising approach for COVID-19 detection. The aim of this study was to develop an AI-assisted diagnostic model that combines chest X-ray images and clinical data to generate a COVID-19 Risk Index (CORI) Score and to implement a deep learning model based on ResNet architecture. Between April 2020 and July 2021, a multicenter cohort study was conducted across three hospitals in Jakarta, Indonesia, involving 367 participants categorized into three groups: 100 COVID-19 positive, 100 with non-COVID-19 pneumonia, and 100 healthy individuals. Clinical parameters (e.g., fever, cough, oxygen saturation) and laboratory findings (e.g., D-dimer and C-reactive protein levels) were collected alongside chest X-ray images. Both the CORI Score and the ResNet model were trained using this integrated dataset. During internal validation, the ResNet model achieved 91% accuracy, 94% sensitivity, and 92% specificity. In external validation, it correctly identified 82 of 100 COVID-19 cases. The combined use of imaging, clinical, and laboratory data yielded an area under the ROC curve of 0.98 and a sensitivity exceeding 95%. The CORI Score demonstrated strong diagnostic performance, with 96.6% accuracy, 98% sensitivity, 95.4% specificity, a 99.5% negative predictive value, and a 91.1% positive predictive value. Despite limitations—including retrospective data collection, inter-hospital variability, and limited external validation—the ResNet-based AI model and the CORI Score show substantial promise as diagnostic tools for COVID-19, with performance comparable to that of experienced thoracic radiologists in Indonesia.

Keywords: COVID-19, diagnostic, scoring system, artificial intelligence, X-ray

Introduction

The coronavirus disease 2019 (COVID-19) pandemic, caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), has highlighted significant global healthcare challenges.



Early and accurate diagnosis of COVID-19 is essential for effective management and mitigation of the disease's spread. The World Health Organization (WHO) has reported that limited diagnostic RT-PCR availability places significant strain on healthcare systems worldwide, highlighting the need for alternative diagnostic tools [1,2]. While RT-PCR remains the gold standard for COVID-19 diagnosis, its limited availability and delayed processing times, especially in remote regions, necessitate complementary diagnostic methods [3,4].

The use of chest X-rays plays a pivotal role in identifying the severity of COVID-19, particularly for moderate to severe cases. Several studies have shown that chest X-rays could detect key COVID-19 imaging patterns, including consolidation (81.3%) [5], reticular interstitial thickening (39.9%) [6], ground-glass opacities (32.5%) [7], nodules (9.3%), and pleural effusion (7.5%) [8]. These abnormalities are typically seen in the lower zones of the lungs. Although CT scans have higher sensitivity (up to 97%) for detecting COVID-19-related lung abnormalities, chest X-rays remain a viable option for triage and monitoring due to their portability and cost-effectiveness [4,9,10]. A study comparing these modalities demonstrated no significant difference in specificity, making chest X-rays particularly valuable in resource-limited settings [11].

Artificial intelligence (AI) has emerged as a transformative tool in medical diagnostics, enabling faster and more accurate disease detection [12]. AI models applied to chest X-rays have achieved diagnostic accuracy rates up to 96%, effectively distinguishing COVID-19 from non-COVID pneumonia and normal cases [13-15]. For instance, Asnaoui *et al.* reported 92.18% accuracy using Inception-ResNetV2 to classify chest X-ray images into normal, bacterial pneumonia, and COVID-19 categories [16].

Despite these advancements, challenges persist. Many AI-driven models depend on imbalanced datasets, lack external validation, or fail to integrate multidimensional clinical and imaging data [14,17-19]. Common limitations include biases related to image acquisition protocols and the lack of generalizability across healthcare settings [18,19]. However, by combining AI with clinical and laboratory data, these challenges could be mitigated to improve diagnostic accuracy and applicability [17]. Therefore, the aim of this study was to address longstanding diagnostic challenges by introducing the CORI Score, a novel diagnostic scoring system that combines imaging analysis with clinical and laboratory data. Additionally, an AI-based ResNet application was designed to support the evaluation and highlight the diagnostic potential of automated image interpretation. By leveraging machine learning models like ResNet, this study sought to advance current diagnostic methodology and provide a scalable solution for healthcare system in low-resource settings. Limitations in prior studies, such as the lack of external validation and challenges in integrating multidimensional data, are also addressed to enhance the generalizability of the proposed approach.

Methods

Study design and patient grouping

A multicenter retrospective was conducted at three general hospitals (Bunda Margonda Central Hospital, Bunda Menteng Central Hospital, and Dr. Cipto Mangunkusumo Central Hospital) in Jakarta between April 2020 and July 2021. A total of 367 patients were admitted during this period. Of these, 267 subjects were utilized for developing the machine learning models, divided into three groups: (1) normal; (2) non-COVID-19 pneumonia; and (3) COVID-19 groups. Additionally, 100 suspected COVID-19 patients were used as an external validation cohort to evaluate the generalizability of the proposed AI models. These 100 patients were not included in the training phase. The discrepancy between the total cohort (367 patients) and the analyzed subset (300 patients) arised from the inclusion criteria specific to the study's modeling phase. The remaining 67 patients were excluded due to incomplete data or poor-quality chest X-rays. This was to ensure that only cases with comprehensive clinical, laboratory, and imaging data were used for analysis and modeling.

Primary and secondary data were collected, encompassing clinical conditions, risk factors (such as contact history with at-risk individuals), comorbidities, and supporting clinical pathology findings. A study flowchart summarizing the methodology is presented in **Figure 1**.

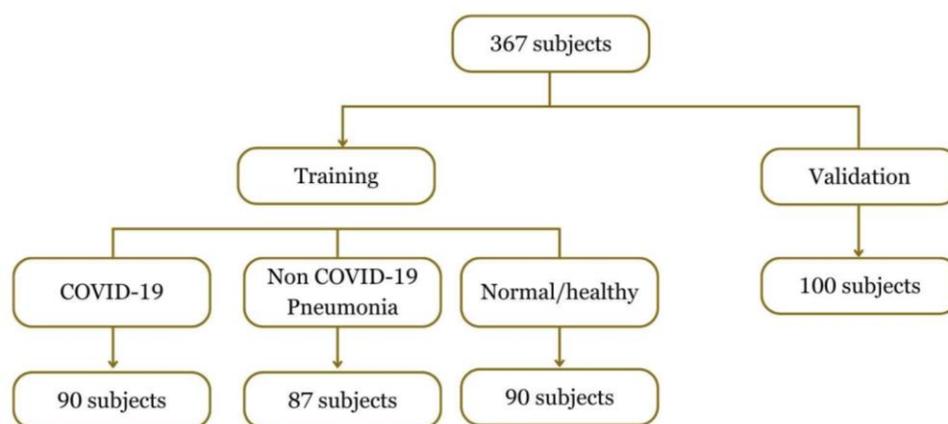


Figure 1. Study flowchart showing the patient groups used in the study.

Patients and criteria

This study included patients aged between 18 and 95 years. Eligible patients had clinical, laboratory, and imaging data available for analysis. For the COVID-19 group, patients were required to have a diagnosis confirmed by RT-PCR testing, which was performed on nasopharyngeal swabs or bronchoalveolar lavage samples in intubated patients. Non-COVID-19 pneumonia patients were those diagnosed clinically and radiologically with pneumonia but confirmed to be negative for COVID-19. The normal group consisted of patients who had no signs of pneumonia or COVID-19, as confirmed by chest X-rays and clinical assessments.

Patients with poor-quality chest X-rays that hindered accurate interpretation were excluded from the study. In addition, patients with incomplete clinical or laboratory data records were excluded to ensure the integrity and completeness of the data used for analysis.

Patient classification

Based on the clinical symptoms, the COVID-19 patients were classified into four categories. Mild COVID-19, patients with headache, fever, respiratory and intestinal symptoms; moderate COVID-19, patients with symptoms of pneumonia with oxygen saturation at rest $>93\%$ and imaging findings of pneumonia signs. Severe COVID-19, patients who meet any of the following criteria: (1) respiratory rate $>30/\text{min}$ (respiratory distress); (2) $\text{PaO}_2/\text{FiO}_2 <300$ mmHg (1 mmHg equal to 0.133 kPa); (3) $<93\%$ oxygen saturation at rest. Very severe COVID-19, patients who meet any of the following criteria: (1) ARDS; (2) required mechanical ventilation; (3) multiple organ failure; (4) shock; (5) intensive care unit is required.

To ensure representativeness, patients were randomly sampled from hospital admissions, with stratification based on clinical characteristics and demographic diversity. This approach minimized selection bias and ensured a balanced distribution of disease severity across the groups.

Clinical and laboratory data collection

Clinical and laboratory data were collected from electronic medical records (EMR) and hospital information systems (HIS) across the three participating hospitals. The clinical data included patient demographics (such as age, sex, and comorbidities), presenting symptoms (e.g., cough, fever, and difficulty breathing), and disease severity based on clinical assessment and imaging results. The severity of COVID-19 was classified according to a set of established criteria, ranging from mild to very severe. Additionally, comorbidities, including hypertension, diabetes mellitus, and cardiovascular disease, were recorded to assess their potential role in disease progression.

For laboratory data, a standardized protocol was followed across the three hospitals to ensure uniformity. The key laboratory parameters collected included peripheral blood counts (e.g., hemoglobin levels and leukocyte count), inflammatory markers such as C-reactive protein (CRP) and D-dimer levels, and serum biochemistry tests, which encompassed electrolytes (e.g., sodium, potassium, and chloride), blood glucose levels, and renal and liver function markers. These tests were standardized through automated systems calibrated to international diagnostic standards. To minimize variability in data collection, teams responsible for collecting clinical and

laboratory data were provided with centralized training, which helped to ensure consistency in the way data were recorded across the different hospitals.

Chest X-ray analysis

Imaging data were obtained using chest X-ray scans performed in either digital anteroposterior or posteroanterior views at the time of hospital admission. Across the three hospitals involved in the study, 67% of the patients underwent anteroposterior imaging, while 33% underwent PA imaging. These imaging procedures were standardized across the facilities to maintain consistency in the diagnostic process and ensure comparability of results.

Six radiologists from three hospitals, who were blinded to the patient groupings, conducted manual evaluations of each lung to ensure unbiased interpretation. Each lung was carefully assessed for abnormalities, including the presence of ground-glass opacities (GGO), consolidation, or other radiological features indicative of pneumonia. The distribution of these abnormalities was also categorized, focusing on whether they were localized to peripheral versus central regions, anterior versus posterior zones, or apical versus basal areas.

Each radiologist evaluated the imaging findings while also considering clinical and laboratory data. To ensure consistency, a joint review discussion was conducted to resolve any discrepancies. Final agreed-upon radiological diagnoses were then used for comparison with AI-based model classification.

This comprehensive radiological evaluation was essential in differentiating normal imaging findings from those associated with non-COVID-19 pneumonia and COVID-19. By systematically analyzing chest X-ray abnormalities alongside clinical and laboratory findings, this study aimed to improve diagnostic accuracy in distinguishing COVID-19 from other conditions.

Sampling and standardization

The sampling method employed for this study involved consecutive sampling for both COVID-19 and non-COVID-19 pneumonia patients, while normal cases were retrospectively identified from the hospital records. This approach ensured that the study sample accurately reflected the population of interest. To ensure consistency across three hospitals, standardized procedures were followed during data collection. This included the use of consistent imaging protocols for chest X-rays, ensuring that all patients' images were taken in a comparable manner. Furthermore, clinical teams across the hospitals utilized a shared data collection template, ensuring that patient information was uniformly recorded. Regular audits were conducted throughout the data collection period to verify the accuracy and completeness of the collected data, minimizing potential sources of bias.

Statistical analysis

Descriptive statistics were first calculated to summarize the demographic and clinical characteristics of the study population. For categorical data, chi-squared tests were used to assess relationships between variables, while continuous variables were analyzed using Mann-Whitney tests to detect any significant differences between groups. To identify the significant predictors for COVID-19 in the study cohort, binary logistic regression was performed. Additionally, receiver operating characteristic (ROC) curves were generated to evaluate the diagnostic performance of the AI model and the CORI Score, which were integral to the study's analysis. The ROC curves helped assess the sensitivity and specificity of these diagnostic tools, providing valuable insights into their effectiveness in distinguishing between normal, non-COVID-19 pneumonia, and COVID-19 cases.

AI model evaluation for COVID-19 classification

A classification approach within supervised learning was utilized in this study to predict class labels from input data. The focus was on classifying COVID-19 through image data analysis using deep learning techniques, as proposed by Asnaoui *et al.* [16] and Wang *et al.* [12].

This study explored various deep learning architectures for image classification in the proposed framework. The models tested included COVID-Net [12], DenseNet [20], ResNet [21], Inception-ResNet [22], DarkCovidNet [23], and CoroNet [17], each known for their effectiveness in computer vision and medical image analysis. These architectures were trained and evaluated

on the study dataset, with a focus on their classification performance for COVID-19 detection. These models were selected due to their high effectiveness in medical image classification, particularly in detecting pneumonia and COVID-19 from chest X-ray images.

In addition to image data, the patient clinical datasets were integrated, combining images, numerical, categorical, and text data. To assess the feasibility of AI-assisted diagnostics, multiple deep learning models were evaluated for their ability to classify chest X-ray images into three distinct categories: normal, non-COVID-19 pneumonia, and COVID-19. Each data type was processed individually, and the features were fused at the end of the model to incorporate comprehensive patient information.

The proposed model with the ResNet convolutional network reached a final accuracy of 91%, a precision of 84%, 94% sensitivity, 92% specificity, and 89% F1-score. ResNet demonstrated the highest classification accuracy, followed by CoroNet, which also exhibited strong performance metrics. A comparison of the performance metrics is detailed in **Table 1**. Following this initial testing, ResNet was selected as the primary model for further evaluation due to its highest classification performance of all tested metrics. ResNet's deep residual connections allows for efficient feature extraction and improved image classification [12,15,21]. A workflow diagram is shown in **Figure 2**.

Table 1. Performance metrics of different models acquired for images

Model	Accuracy	Sensitivity	Specificity	Precision	F1-Score
ResNet	0.91	0.94	0.92	0.84	0.89
DenseNet	0.87	0.88	0.89	0.74	0.78
InceptionResNet	0.87	0.94	0.86	0.85	0.73
DarkCovidNet	0.91	0.76	0.89	0.73	0.69
CoroNet	0.91	0.94	0.92	0.87	0.81
COVID-Net	0.31	1.00	0.00	0.31	0.48

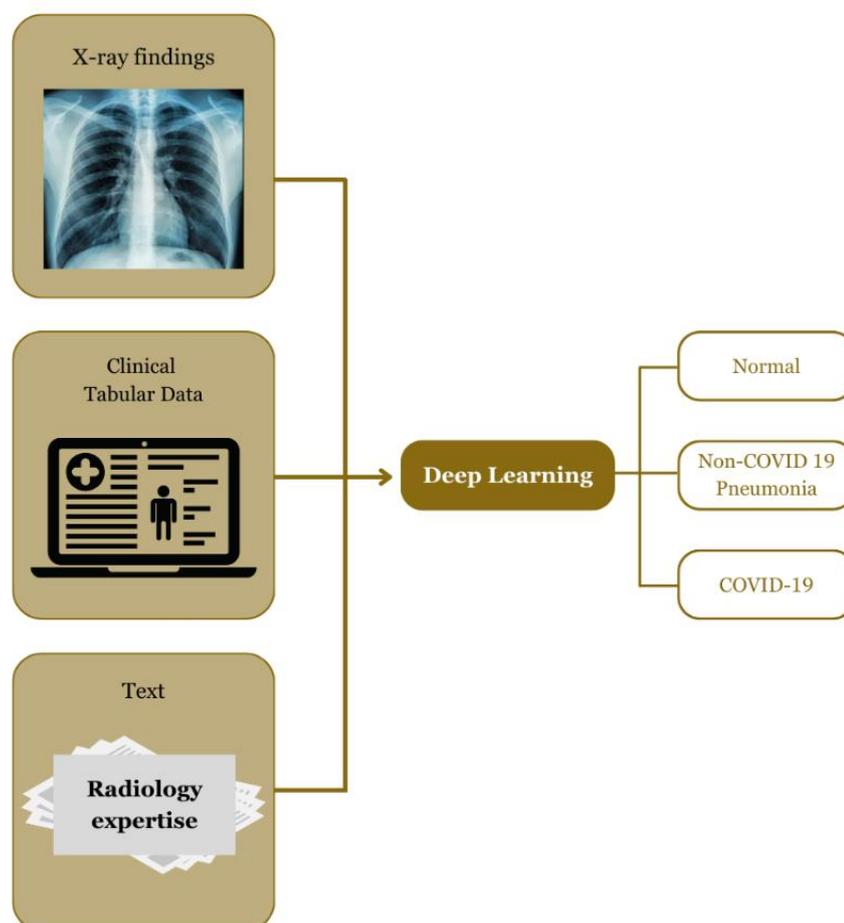


Figure 2. Image-patient data model architecture with ResNet.

Deployment and external validation of ResNet model using Kalamakara-AI database

Lastly to further validate our AI model, the ResNet model was integrated into Kalamakara-AI, a cloud-based web platform designed for easy use and sharing among medical professionals. This AI system is a browser-based web application (<http://kalamakara-ai.ui.ac.id/>) that takes chest X-ray images and prompts users to anonymously fill out a form for the patient's clinical examination, including temperature, fever duration, other conditions, and comorbidities. Some examples of COVID-19 symptoms included in the form are cough, flu, sore throat, headache, faintness, shivering, nausea, diarrhea, muscle ache, abdominal pain, and breathing difficulty. The images were classified into normal imaging, non-COVID-19 pneumonia imaging, or COVID-19 imaging. The system outputted the probability score that predicted the final diagnosis of the patients. The Kalamakara-AI web application could predict COVID-19 in just a few seconds with a 100% probability score and outputs a 0% probability for pneumonia and normal cases, respectively. This platform allowed the AI system to be tested on real-world clinical data, where healthcare professionals could upload chest X-ray images and receive automated diagnostic predictions.

Development of the CORI Score

To enhance diagnostic accuracy, the COVID-19 Risk Index (CORI) Score was designed to integrate clinical, laboratory, and imaging data into a structured model. By incorporating these multiple parameters, the CORI score aimed to provide an accessible, structure tool for clinical decision-making without relying on complex computational models. The key steps in this process, from data collection to final classification into normal, COVID-19, or non-COVID-19 pneumonia, ensuring a more comprehensive evaluation are outlined in **Figure 3**.

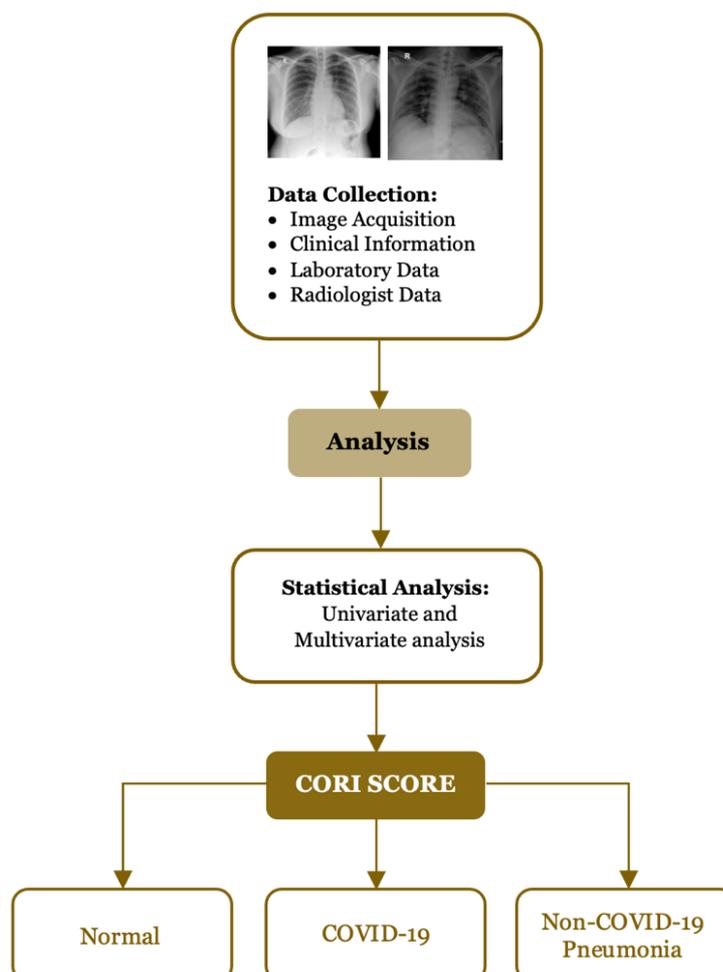


Figure 3. Model architecture to build CORI score.

Although AI-assisted imaging analysis was explored in this study, the CORI score was developed independently of deep learning model. The AI model functioned exclusively as standalone classifiers for chest X-ray interpretation and was not incorporated into the CORI score's calculation. Instead, the CORI score relied on logistic regression modeling, where each clinical, laboratory, and imaging parameter was assigned a weighted coefficient based on statistical significance.

Results

Patients' characteristics

Of the initial 367 patients, 67 were excluded due to incomplete data or poor-quality chest X-ray images, leaving 300 patients who met the inclusion criteria for analysis. A total of 300 patients then included in the study consisted 100 patients with normal chest imaging, 100 non-COVID-19 pneumonia patients, and 100 RT-PCR-confirmed COVID-19 cases. The characteristics of patients included in the study are presented in **Table 2**. The study sample included 151 men (50.3%) and 149 women (49.7%), with a mean age of 70 years (range: 36–95 years). There was a significant difference in age distribution ($p=0.004$), while gender distribution was not statistically significant ($p=0.581$).

In terms of disease severity, the majority of the COVID-19 group had moderate pneumonia (47%), while non-COVID-19 pneumonia patients had mild pneumonia (59%), and the normal imaging group was predominantly mild (94%) ($p<0.001$) (**Table 2**). Symptoms at presentation also differed between groups. Cough was the most frequent symptom in COVID-19 patients (76%), while 50% of non-COVID-19 pneumonia patients also reported cough. In the normal imaging group, malaise was the most commonly reported symptom (13%). Although a significant p -value (<0.001) was observed for disease severity across the groups, no significant differences were found for symptoms like malaise and shivering (**Table 2**).

Of the COVID-19 patients, 48% had underlying diseases, with hypertension ($n=14$) and diabetes mellitus ($n=14$) being the most common comorbidities (**Table 2**). In the non-COVID-19 pneumonia group, 47% had hypertension ($n=16$), and 9% had malignancy. The normal imaging group also had patients with comorbidities such as hypertension and malignancy.

Table 2. Characteristics of subjects included in the study

Characteristic	Subject group			p-value
	COVID-19 (n=100)	Non COVID-19 (n=100)	Normal (n=100)	
Age (years)				0.004
18–25	7 (7%)	3 (3%)	5 (5%)	
26–35	21 (21%)	14 (14%)	25 (25%)	
36–45	30 (30%)	17 (17%)	26 (26%)	
46–55	18 (18%)	15 (15%)	20 (20%)	
56–65	14 (14%)	24 (24%)	11 (11%)	
>65	10 (10%)	27 (27%)	13 (13%)	
Sex				0.581
Male	53 (53%)	51 (51%)	46 (46%)	
Female	47 (47%)	48 (48%)	54 (54%)	
Symptom				
Fever	69 (69%)	18 (18%)	10 (10%)	0.001
Cough	76 (76%)	50 (50%)	7 (7%)	<0.001
Anosmia	18 (18%)	1 (1%)	0 (0%)	<0.001
Ageusia	4 (4%)	0 (0%)	0 (0%)	0.017
Malaise	24 (24%)	23 (23%)	13 (13%)	0.099
Flu	23 (23%)	2 (2%)	3 (3%)	<0.001
Muscle pain	17 (17%)	0 (0%)	2 (2%)	<0.001
Nausea vomitus	20 (20%)	3 (3%)	9 (9%)	<0.001
Hard breath	21 (21%)	6 (6%)	5 (5%)	<0.001
Abdominal pain	32 (32%)	36 (36%)	6 (6%)	<0.001
Shiver	2 (2%)	4 (4%)	1 (1%)	0.359
Diarrhea	11 (11%)	0 (0%)	3 (3%)	0.001
Headache	20 (20%)	3 (3%)	4 (4%)	<0.001
Comorbidity				
Obesity	2 (2%)	0 (0%)	1 (1%)	0.364

Characteristic	Subject group			p-value
	COVID-19 (n=100)	Non COVID-19 (n=100)	Normal (n=100)	
Hypertension	14 (14%)	16 (16%)	7 (7%)	0.127
Diabetes	14 (14%)	6 (6%)	2 (2%)	0.004
Chronic kidney disease	2 (2%)	1 (1%)	2 (2%)	0.816
STEMI	2 (2%)	1 (1%)	0 (0%)	0.364
NSTEMI	1 (1%)	0 (0%)	1 (1%)	0.604
COPD	0 (0%)	2 (2%)	0 (0%)	0.134
Tuberculosis	2 (2%)	8 (8%)	1 (1%)	0.017
Autoimmunity	1 (1%)	3 (3%)	1 (1%)	0.443
Malignancy	4 (4%)	10 (10%)	9 (9%)	0.232
Hepatitis	4 (4%)	0 (0%)	1 (1%)	0.071
HIV	2 (2%)	0 (0%)	2 (2%)	0.363
Laboratory				
Hemoglobin, median	11.9 g/dL	11.9 g/dL	12.75 g/dL	0.104
Leukocyte, median	8865 cells/ μ L	9000 cells/ μ L	8220 cells/ μ L	0.292
D-dimer, median	3270 ng/mL	3270 ng/mL	1485 ng/mL	0.004
CRP, median	37.8 mg/L	37.8 mg/L	22.9 mg/L	0.376
Blood glucose level, median	108 mg/dL	109.5 mg/dL	103 mg/dL	0.134
Aspartate aminotransferase, median	24 U/L	24 U/L	21 U/L	0.098
Alanine aminotransferase, median	23.5 U/L	23.5 U/L	21 U/L	0.119
Sodium, median	137 mmol/L	137 mmol/L	137 mmol/L	0.957
Potassium, median	3.9 mmol/L	3.9 mmol/L	3.9 mmol/L	0.800
Chloride, median	99 mmol/L	99 mmol/L	100.8 mmol/L	0.344
Urea, median	28.85 mg/dL	28.85 mg/dL	25.45 mg/dL	0.490
Creatinine, median	28.85 mg/dL	0.9 mg/dL	0.8 mg/dL	0.167
Imaging analysis				
Apical vs basal				<0.001
Apical	20 (20%)	32 (32%)		
Basal	80 (80%)	44 (44%)		
Apical and basal	0 (0%)	24 (24%)		
Anterior vs posterior				<0.001
Anterior	91 (91%)	53 (53%)		
Posterior	2 (2%)	5 (5%)		
Anterior and posterior	7 (7%)	42 (42%)		
Unilateral vs bilateral				0.138
Unilateral	25 (25%)	33 (33%)		
Bilateral	75 (75%)	67 (67%)		
Peripheral vs central				0.001
Peripheral	18 (18%)	8 (8%)		
Central	56 (56%)	42 (42%)		
Peripheral and central	26 (26%)	50 (50%)		
Opacity vs non-opacity				0.217
Opacity	39 (39%)	51 (51%)		
Non-opacity	48 (48%)	40 (40%)		
Opacity and non-opacity	13 (13%)	9 (9%)		
Consolidation vs non-consolidation				0.001
Consolidation	12 (12%)	13 (13%)		
Non-consolidation	88 (88%)	87 (87%)		
Ground-glass opacities vs non-ground-glass opacities				<0.001
Ground-glass opacities	25 (25%)	16 (16%)		
Non-ground-glass opacities	75 (75%)	84 (84%)		

Findings from Kalamakara-AI external evaluation

Out of 100 total cases uploaded into the Kalamakara-AI system, the model predicted COVID-19 in 39 patients (39%), non-COVID-19 pneumonia in 35 patients (35%), and identified normal chest imaging in 22 patients (22%), all with a prediction confidence score of 80% or higher. The remaining 4% of cases yielded prediction confidence below 80% and were therefore not conclusively classified by the AI. These results were generated through an integrated assessment combining chest X-ray image analysis with clinical information provided by the attending physicians.

Building CORI score

The CORI score (COVID-19 Risk Index) was developed using a logistic regression model, designed to assess the likelihood of a patient having COVID-19 based on key clinical symptoms,

laboratory, and radiological findings interpreted by radiologist. These predictors were identified through univariate analysis, which revealed significant associations with COVID-19 infection. Binary logistic regression showed that demographics, such as age group (e.g., 36–45 years), were significant predictors, with older age groups generally indicating higher risk. Key symptoms, including cough, fever, flu-like symptoms, and anosmia, contributed a score based on their presence and severity. Laboratory markers, like AST, ALT, and creatinine levels, reflected the patient's health and risk of complications. Imaging findings from chest X-rays, including the distribution and type of abnormalities (e.g., ground-glass vs non-ground-glass opacities), were also essential for the model. Each factor was attributed a proportional value, generating a cumulative risk score, where higher scores indicated greater likelihood of COVID-19. The model structure is illustrated in **Figure 3**.

Significant predictors of COVID-19 case identification

The binary logistic regression model identified significant predictors for the identification of COVID-19 cases and the results are presented in **Table 3**. The binary logistic regression analysis revealed several key factors significantly associated with COVID-19. These findings provided a robust foundation for the development of the CORI Score by integrating clinical, imaging, and laboratory predictors. The strongest predictor of COVID-19 was group classification ($\beta_1=6.633$, $p<0.001$), as anticipated, given the model's aim to differentiate COVID-19 cases from non-COVID-19 pneumonia and normal groups. The age group of 36–45 years showed significant disease correlation, as indicated in the logistic regression model. This group was more likely to be affected by COVID-19 compared to others, aligning with observed patterns in the demographic data. Chest X-ray findings played a crucial role in distinguishing COVID-19. Basal abnormalities (Apical vs Basal: $\beta_3=0.955$, $p<0.001$) were strongly associated with COVID-19, reflecting known imaging patterns of the disease. The presence of opacities (opacity vs non-opacity: $\beta_4=0.958$, $p<0.001$) was also highly predictive. Conversely, the absence of ground-glass opacities (GGO vs Non-GGO: $\beta_5=-1.048$, $p=0.002$) significantly reduced the likelihood of COVID-19, highlighting the distinct radiological features often observed in these patients. Certain clinical symptoms were identified as significant predictors. Fever ($\beta_7=0.444$, $p=0.020$) and flu-like symptoms ($\beta_8=0.789$, $p<0.001$) showed strong positive associations with COVID-19, reaffirming their diagnostic relevance. Interestingly, cough ($\beta_6=-0.394$, $p=0.048$) was less predictive, possibly due to its nonspecific nature and presence in various other respiratory conditions. Among laboratory findings, elevated AST levels ($\beta_{11}=0.022$, $p=0.045$) were found to be mildly predictive of COVID-19. Although its effect size was modest, this underscores the importance of incorporating laboratory markers for a more comprehensive diagnostic model.

Table 3. Binary logistic regression model for significant predictor of COVID-19 case identification

Variable	Coefficient (β)	p-value
Group study	6.633	<0.001
Anterior vs posterior	-0.106	0.035
Apical vs basal	0.955	<0.001
Opacity vs non-opacity	0.958	<0.001
Ground-glass opacities vs non-ground-glass opacities	-1.048	0.002
Presence of cough	-0.394	0.048
Presence of fever	0.444	0.020
Presence of flu-like symptom	0.789	<0.001
Presence of anosmia	-0.523	0.030
Serum creatinine level	-0.009	0.080
Aspartate aminotransferase level	0.022	0.045
Alanine aminotransferase level	-0.200	0.058

The CORI Score was calculated by applying these coefficients to individual patient data, with a higher score indicating a greater likelihood of COVID-19. The diagnostic performance of the CORI Score was evaluated using ROC curve analysis. The ROC curves demonstrated a significant area under the curve for diagnosing COVID-19, indicating the model's robustness in distinguishing between COVID-19 and other conditions.

Diagnostic performance of the CORI score

The diagnostic test of CORI Score demonstrated high performance, with a sensitivity of 98%, specificity of 95.4%, and overall accuracy of 96.6%. The negative predictive value (NPV) was 99.5%, while the positive predictive value (PPV) was 91.08%. The area under the curve (AUC) was 0.982 (95% confidence interval: 0.965–0.999), indicating excellent discriminative ability ($p < 0.001$) (Figure 4).

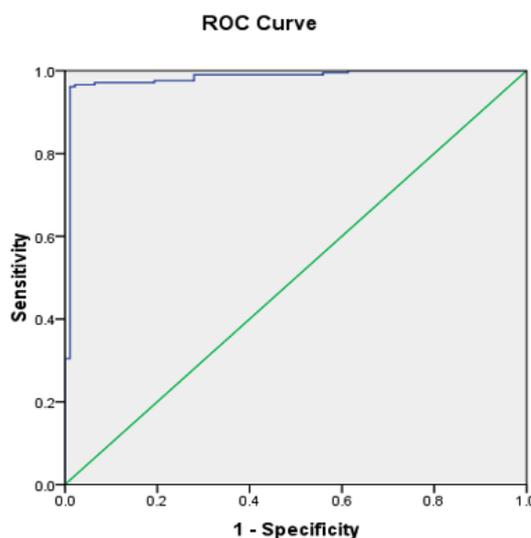


Figure 4. ROC of CORI score showing an area under the curve of 0.982 (95% confidence interval (CI): 0.965–0.999) and $p < 0.001$ for diagnosing COVID-19.

Testing and validation

The CORI score's diagnostic accuracy was further validated with a hold-out test set of 300 patients, where structured clinical, laboratory, and chest X-ray findings were collected. These patients were divided as follows: 100 images from COVID-19 patients, 100 images from non-COVID-19 pneumonia patients, and 100 images from normal patients. The ROC curves of the CORI Score displayed a high AUC for diagnosing COVID-19, confirming its high diagnostic accuracy.

In comparison, the ResNet model, a deep learning algorithm used for image classification, was also evaluated on this dataset. Unlike the CORI score, ResNet was directly applied to the raw chest X-ray images in combination with clinical parameters. It achieved an overall accuracy of 82%, correctly identifying 82 out of 100 COVID-19 cases. It demonstrated a sensitivity of 100%, detecting all COVID-19 cases, but had a low specificity of 18%, misclassifying a significant number of non-COVID-19 pneumonia and normal cases as COVID-19. While ResNet performed well in detecting pneumonia cases, its limited specificity resulted in a higher rate of false positives for COVID-19.

The CORI score demonstrated strong diagnostic performance by incorporating a combination of clinical, laboratory, and interpreted radiological data, making it a highly effective tool for diagnosing COVID-19 cases. During binary logistic regression, COVID-19 infection showed a significant relation with the identified predictors, contributing to the model's high sensitivity and specificity. The score's ROC curve analysis highlighted its superior diagnostic performance, with a sensitivity of 98%, specificity of 95.4%, and accuracy of 96.6%.

However, further prospective validation is necessary to determine the optimal cut-off point for clinical use. Establishing a validated cut-off is essential to maximize diagnostic accuracy while minimizing false positives and false negatives, ensuring the CORI score can be effectively applied across diverse patient populations. Without this validation, it remains a promising yet incomplete diagnostic tool, requiring additional studies to refine its clinical applicability, threshold values, and real-world performance.

Unlike standalone AI-based models such as ResNet, which rely solely on raw radiological images and clinical features processed through complex AI algorithms, the CORI Score offers a

more interpretable, structured approach by integrating explicitly identified clinical parameters. Rather than comparing them head-to-head, their outcomes were observed in parallel to evaluate whether AI-generated results could reflect real-world diagnostic logic embedded within CORI scoring system. This makes the CORI score more comprehensive tool for early detection and management of COVID-19, particularly in settings where AI-based diagnostics may be limited by computational or infrastructure constraints.

Interobserver agreement

Interobserver agreement in radiological assessment plays a critical role in validating the reliability and reproducibility of imaging-based diagnostics. In this study, six radiologists from three hospitals independently evaluated the chest X-ray dataset, interpreting the images without prior knowledge of the AI model classification results. Each of them assessed the images while also considering the corresponding clinical symptoms and laboratory findings. The individual interpretations were then reviewed collectively in a joint discussion, confirming that potential discrepancies were addressed before finalizing the radiological diagnoses.

Following this independent assessment, the final agreed-upon radiological interpretations were compared to the AI-based classification results. The findings demonstrated a high level of concordance between the radiologist's interpretations and the Kalamakara-AI outputs, indicating that it effectively captured key radiological features relevant to COVID-19 classification. This agreement supports the validity of imaging data used in this study, proving the variability in radiological assessments did not introduce bias into the development of the CORI score.

Discussion

Early and accurate detection and diagnosis of COVID-19 remains a crucial challenge, particularly in resource-limited settings where RT-PCR testing is not always accessible [24,25]. Prior research has indicated the effectiveness of chest X-ray images in diagnosing COVID-19 [26,27]. This study presented an AI-driven system based on ResNet, designed to detect the presence of COVID-19 in chest X-ray images and differentiate it from normal/healthy conditions or non-COVID-19 pneumonia [17]. The CORI score, developed in this study, used logistic regression modeling to integrate clinical symptoms, laboratory markers, and radiological findings into a structured diagnostic model, offering a practical alternative to traditional methods. Elevated D-dimer levels, a key finding in this study, are commonly observed in COVID-19 patients [28], indicating infection severity, immune status, and worse prognosis, especially in cases with disseminated intravascular coagulation due to fibrin breakdown in blood clots [29,30]. ROC curve analysis confirmed its strong diagnostic performance, with a sensitivity of 98%, specificity of 95.4%, and overall accuracy of 96.6%, underlining its potential for aiding clinical decision-making.

In a previous study, combining radiological imaging and clinical data with AI algorithms was reported to be more accurate in detecting COVID-19 infection [19]. AI-assisted image analysis was explored in this study using Kalamakara-AI web application, with ResNet achieving the highest accuracy (91%), showcasing its potential for automated COVID-19 screening. Therefore, it has been ensured that the Kalamakarata-AI-based web applications is publicly accessible, allowing healthcare professionals to test its implementation in clinical practice. Critical feedback from medical professionals will provide additional guidance to improve the Kalamakara-AI model for a better one. Prior research has shown that deep learning models, particularly convolutional neural networks (CNNs), can perform comparably to radiologists in detecting pneumonia and COVID-19-related abnormalities [22]. However, AI models often require large datasets and specialized computational resources, limiting their feasibility in low-resource settings. Unlike AI-based models that rely solely on imaging, the CORI score integrates multiple diagnostic modalities, making it more accessible for diverse healthcare settings. By combining different diagnostic parameters, the CORI score improves predictive accuracy while reducing dependence on advanced imaging technology, offering a practical alternative where AI-based diagnostics like Kalamakara-AI may not be available.

Despite its promising diagnostic performance, the CORI score requires further validation before widespread clinical implementation. This study has several limitations that must be addressed in future research. A prospective validation study is needed to assess the CORI score's

performance in real-world settings. While the CORI score showed high sensitivity and specificity, an optimal diagnostic threshold has not been formally established. Further multi-center validation studies are necessary to define a clinically relevant cut-off point that minimizes false positives and false negatives, assuring consistent diagnostic accuracy across different populations. Variability in clinical and radiological data remains a factor, and although six radiologists agreed with the AI-based imaging results, radiological assessment can still be subject to interobserver variability in clinical practice. Additional studies should explore whether standardized radiological scoring criteria can further improve the reproducibility of imaging-based diagnostics.

This study was conducted in a limited number of hospitals, and the CORI score's applicability to other geographic regions and patient demographics remains unknown. Further validation across diverse healthcare settings is essential to confirm its generalizability and robustness. Additionally, although the CORI score was shown to be effective, its performance should be compared with existing clinical scoring systems, such as qSOFA, to determine its relative effectiveness in predicting COVID-19 severity and progression [31]. These refinements will be crucial in optimizing the CORI score as a practical, evidence-based diagnostic tool for COVID-19 and future respiratory infections.

Conclusion

The AI model and a new diagnostic score have been developed in Indonesia to identify COVID-19 infection utilizing chest X-ray combined clinical and laboratory data. The ResNet exhibited the highest accuracy, sensitivity, and specificity relative to other models. In the diagnosis of COVID-19, the AI model and CORI Score proved promising potential and performed in a manner comparable to that of experienced thoracic radiologists in Indonesia. After validation, the AI Model and CORI Score are appropriate for wide use in Indonesia.

Ethics approval

This study has been approved by the Ethics Committee of the Faculty of Medicine, University of Indonesia, Jakarta, Indonesia (Protocol Number: 20-05-0509). All subjects provided written informed consent before participating in the study.

Acknowledgments

None.

Competing interests

All the authors declare that there are no conflicts of interest.

Funding

This research was funded by Direktorat Inovasi & Science Tecno Park (Techno Park Innovation and Science Directory), Universitas Indonesia.

Underlying data

Derived data supporting the findings of this study are available from the corresponding author on request.

Declaration of artificial intelligence use

This study used artificial intelligence (AI) tools and methodologies in the following capacities. Data analysis and modeling: multiple deep learning architectures were utilized for image classification and prediction tasks, including COVID-Net, Dense-Net, ResNet, Inception-ResNet, DarkCovidNet, CoroNet. Among these, ResNet demonstrated the highest accuracy and was selected for further model integration. These models were implemented using Python and TensorFlow, ensuring robust deep learning capabilities for processing chest X-ray images. Data preprocessing: AI-assisted techniques, such as image augmentation, feature scaling, and word embedding (Word2Vec for text processing), were applied to enhance dataset quality and optimize

model training. Visualization: AI tools, including Matplotlib, TensorFlow visualization libraries, and cloud-based web deployment (<http://kalamakara-ai.ui.ac.id/>), were used for generating model performance graphs, ROC curves, and user-interface displays. We confirm that all AI-assisted processes were critically reviewed by the authors to ensure the integrity and reliability of the results. The final decisions and interpretations presented in this article were solely made by the authors.

How to cite

Kamelia T, Zulkarnaen B, Septiyanti W, *et al.* Designing the CORI score for COVID-19 diagnosis in parallel with deep learning-based imaging models. *Narra J* 2025; 5 (2): e1606 - <http://doi.org/10.52225/narra.v5i2.1606>.

References

1. World Health Organization. Coronavirus disease (COVID-19). Geneva: World Health Organization; 2021.
2. Hargreaves J, Davey C, Hargreaves J, *et al.* Three lessons for the COVID-19 response from pandemic HIV. *Lancet HIV* 2020;7(5):e309-e311.
3. Quraishi E, Jibuaku C, Lisik D, *et al.* Comparison of clinician diagnosis of COVID-19 with real time polymerase chain reaction in an adult-representative population in Sweden. *Respir Res* 2023;24(1):10.
4. Borakati A, Perera A, Johnson J, *et al.* Diagnostic accuracy of X-ray versus CT in COVID-19: A propensity-matched database study. *BMJ Open* 2020;10(11):e042946.
5. Yasin R, Gouda W. Chest X-ray findings monitoring COVID-19 disease course and severity. *Egypt J Radiol Nucl Med* 2020;51(1):193.
6. Abougazia A, Alnuaimi A, Mahran A, *et al.* Chest X-ray findings in COVID-19 patients presenting to primary care during the peak of the first wave of the pandemic in Qatar: Their association with clinical and laboratory findings. *Pulm Med* 2021;2021:4496488.
7. El Houby EMF. COVID-19 detection from chest X-ray images using transfer learning. *Sci Rep* 2024;14(1):11639.
8. Pezzutti DL, Makary MS. Role of imaging in diagnosis and management of COVID-19: Evidence-based approaches. In: Rezaei N, editor. *The COVID-19 aftermath*. Cham: Springer International Publishing; 2024.
9. Tenda ED, Yulianti M, Asaf MM, *et al.* The importance of chest CT scan in COVID-19. *Acta Med Indones* 2020;52(1):68-73.
10. Mahdavi Sharif P, Nematizadeh M, Saghadzadeh M, *et al.* Computed tomography scan in COVID-19: A systematic review and meta-analysis. *Pol J Radiol* 2022;87:e1-e23.
11. Rousan LA, Elobeid E, Karrar M, *et al.* Chest X-ray findings and temporal lung changes in patients with COVID-19 pneumonia. *BMC Pulm Med* 2020;20(1):245.
12. Wang L, Lin ZQ, Wong A. COVID-Net: A tailored deep convolutional neural network design for detection of COVID-19 cases from chest X-ray images. *Sci Rep* 2020;10(1):19549.
13. Arora N, Banerjee AK, Narasu ML. The role of artificial intelligence in tackling COVID-19. *Future Virol* 2020;15(11):717-724.
14. Esposito A, Casiraghi E, Chiaraviglio F, *et al.* Artificial intelligence in predicting clinical outcome in COVID-19 patients from clinical, biochemical and a qualitative chest X-ray scoring system. *Rep Med Imaging* 2021;14:27-39.
15. Zhu J, Shen B, Abbasi A, *et al.* Deep transfer learning artificial intelligence accurately stages COVID-19 lung disease severity on portable chest radiographs. *PLoS One* 2020;15(7):e0236621.
16. El Asnaoui K, Chawki Y. Using X-ray images and deep learning for automated detection of coronavirus disease. *J Biomol Struct Dyn* 2021;39(10):3615-3626.
17. Shamout FE, Shen Y, Wu N, *et al.* An artificial intelligence system for predicting the deterioration of COVID-19 patients in the emergency department. *Npj Digit Med* 2021;4(1):80.
18. Tsiknakis N, Trivizakis E, Vassalou E, *et al.* Interpretable artificial intelligence framework for COVID-19 screening on chest X-rays. *Exp Ther Med* 2020;20(2):727-735.
19. Mei X, Lee HC, Diao KY, *et al.* Artificial intelligence-enabled rapid diagnosis of patients with COVID-19. *Nat Med* 2020;26(8):1224-1228.
20. Müller AC, Guido S. *Introduction to machine learning with Python: A guide for data scientists*. First edition. Sebastopol, CA: O'Reilly Media, Inc; 2016.
21. Srinivasan S, editor. *Guide to big data applications*. Cham: Springer International Publishing; 2018.

22. Nishio M, Noguchi S, Matsuo H, *et al.* Automatic classification between COVID-19 pneumonia, non-COVID-19 pneumonia, and the healthy on chest X-ray image: Combination of data augmentation methods. *Sci Rep* 2020;10(1):17532.
23. Ronneberger O, Fischer P, Brox T. U-Net: Convolutional networks for biomedical image segmentation. In: Navab N, Hornegger J, Wells WM, Frangi AF, editors. *Medical image computing and computer-assisted intervention – MICCAI 2015*. Cham: Springer International Publishing; 2015.
24. Kim JK, Crimmins EM. How does age affect personal and social reactions to COVID-19: Results from the national understanding America study. *PLoS One* 2020;15(11):e0241950.
25. Cavallo AU, Troisi J, Forcina M, *et al.* Texture analysis in the evaluation of COVID-19 pneumonia in Chest X-ray images: A proof of concept study. *Curr Med Imaging Former Curr Med Imaging Rev* 2021;17(9):1094-1102.
26. Shi H, Han X, Jiang N, *et al.* Radiological findings from 81 patients with COVID-19 pneumonia in Wuhan, China: A descriptive study. *Lancet Infect Dis* 2020;20(4):425-434.
27. Kong W, Agarwal PP. Chest imaging appearance of COVID-19 infection. *Radiol Cardiothorac Imaging* 2020;2(1):e200028.
28. Ibrahim ME, AL-Aklobi OS, Abomughaid MM, *et al.* Epidemiological, clinical, and laboratory findings for patients of different age groups with confirmed coronavirus disease 2019 (COVID-19) in a hospital in Saudi Arabia. *PLoS One* 2021;16(4):e0250955.
29. Jin YH, Cai L, Cheng ZS, *et al.* A rapid advice guideline for the diagnosis and treatment of 2019 novel coronavirus (2019-nCoV) infected pneumonia (standard version). *Mil Med Res* 2020;7(1):4.
30. Statsenko Y, Al Zahmi F, Habuza T, *et al.* Prediction of COVID-19 severity using laboratory findings on admission: Informative values, thresholds, ML model performance. *BMJ Open* 2021;11(2):e044500.
31. Liu S, Yao N, Qiu Y, *et al.* Predictive performance of SOFA and qSOFA for in-hospital mortality in severe novel coronavirus disease. *Am J Emerg Med* 2020;38(10):2074-2080.