P-ISSN 2586-9000
E-ISSN 2586-9027
Homepage : https://tci-thaijo.org/index.php/SciTechAsia
Science & Technology Asia
Vol.28 No.4 October - December 2023
Page: [305-313]

Original research article

Validation of a Deep Learning-Based Automated Detection Algorithm for Mobile Chest Radiographs of Patients with SARS-CoV-2 Infection in a Field Hospital

Warit Tarathipmon*, Amolchaya Kwankua, Utairat Chaumrattanakul, Pisit Wattanaruangkowit

Department of Radiology, Faculty of Medicine, Thammasat University, Pathum Thani 12120, Thailand

Received 3 June 2023; Received in revised form 10 November 2023
Accepted 7 December 2023; Available online 27 December 2023

ABSTRACT

Due to the rapid spread of SARS-CoV2 infection, many hospitals had an influx patients along with more chest radiographs requiring interpretation. The deep learning-based automatic detection (DLAD) algorithm was applied to help physicians and radiologists in reading this overwhelming number of chest radiographs. The aim of this study was to validate the diagnostic performance in COVID-19 pneumonia detection of the DLAD for mobile chest radiographs in patients with SARS-CoV-2 infection. Chest radiographs of patients with RT-PCR confirmed SARS-CoV-2 infection were included. The diagnostic performance of DLAD in the detection of COVID-19 pneumonia with mobile chest radiographs was evaluated in comparison to the prior interpretation by the radiologist. The sensitivity and specificity of DLAD in identification of pneumonia were 27.6% (95%CI 22.3%-33.4%) and 99.8% (95%CI 98.6%-100%) respectively. PPV and NPV were 98.7% (95%CI 92.8%-100%) and 67.4% (95%CI 63.5%-71.2%) respectively, with an AUC of 0.64 (95%CI 0.61-0.66). The duration from the onset of symptoms to the time of chest radiography was a significant predictor of pneumonia (P=0.013) whereas sex, age, BMI, and symptoms were not. DLAD is a tool with excellent specificity and PPV for COVID-19 pneumonia detection but has low sensitivity and AUC.

Keywords: Chest radiographs; COVID-19 pneumonia; Deep learning-based automatic detection (DLAD) algorithm; SARS-CoV2 infection

1. Introduction

The chest radiograph is commonly used worldwide as a screening tool or for obtaining diagnosis for pulmonary parenchymal, airway, mediastinal, cardiac, pleural, and chest wall abnormalities [1]. Its main limitation is the

***Corresponding author:** warit_t@tu.ac.th
doi: 10.14456/scitechasia.2023.90

potential for diagnostic errors, especially when interpreted by less-experienced interpreters.

The deep learning-based automatic detection (DLAD) algorithm was developed to help obtain fast patient triage, improved interpretation performance, and decreased workload. In our institution, the DLAD is installed into the mobile radiography unit. After a chest radiograph is done, the DLAD generates the location of detected lesions, the abnormality score reflecting the probability of the existence of detected lesions, and a summary of the findings.

In patients with the novel coronavirus severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection, chest radiography and computed tomography (CT) play an important role in the detection of viral pneumonia and in follow-up evaluation [2]. Due to the increase in patients admitted to hospitals, there was a large number of chest radiographs requiring interpretation. The DLAD with the mobile chest radiograph was applied to help physicians and radiologists in interpreting this overwhelming number of chest radiographs.

The primary objective of this study was to validate the diagnostic performance in COVID-19 pneumonia detection of the DLAD for mobile chest radiographs in patients with SARS-CoV-2 infection admitted into the Thammasat University Field Hospital.

2. Materials and Methods

This study was approved by the Human Research Ethics Committee of Thammasat University (Medicine). The number of Certificate of Approval is 319/2021. This single-center study was a diagnostic cross-sectional study with a retrospective approach.

2.1 Participants

Included in this study were COVID-19 patients who underwent mobile chest radiography with DLAD algorithm at the Thammasat University Field Hospital from 1 May 2021 to 31 July 2021. Criteria for admission into the Thammasat University Field Hospital are as follows: 1) patients with mild symptoms and low risk, 2) patients who had been treated in any hospital for at least 48 hours with stable vital signs, 3) asymptomatic patients with no risk factors, capable of self-care, able to communicate well, and have no risk of psychiatric disorders nor aggressive behavior. However, patients who did not fulfill the above criteria might be considered for admission into the Thammasat University Field Hospital upon the approval by the infectious specialist. In this study, participants were eligible to be included if their age fell between 18 to 70 years and were confirmed to have SARS-CoV-2 infection by real-time reverse-transcriptase polymerase chain reaction (RT-PCR). The younger and older patients were not included in this study due to anatomical reasons and age-related changes in the lungs of the elderly. Patients were excluded if they had any chest devices, a history of thoracic surgery or breast augmentation, known thoracic neoplasms, or severe artifacts on the image. Age, sex, body mass index (BMI), onset of symptoms, and date of chest radiography were recorded. Only the first chest radiograph was used for analysis.

2.2 Deep learning-based automatic detection (DLAD) algorithm

The DLAD algorithm (Lunit INSIGHT, Lunit Inc.) assessed in our study was installed on the mobile chest radiography unit. This DLAD was approved by the Ministry of Food and Drug Safety of Korea. The algorithm was developed by a deep convolutional neural network with dense blocks and originally trained with 54,221 normal chest radiographs and 35,613 abnormal chest radiographs in four major diseases including pulmonary malignancy, pneumonia, pulmonary tuberculosis, and pneumothorax [3, 4].

This algorithm can detect ten radiologic findings which are consolidation, pleural effusion, pneumothorax, atelectasis, cardiomegaly, fibrosis, nodule, pneumoperitoneum, calcification, and mediastinal widening. In this study, we focused on the detection of consolidations representing pneumonia. However, other findings were also recorded. The DLAD revealed an abnormality score between 0% to 100% reflecting the probability of the existence of detected lesions. The report of the lesion and its location were displayed when the probability score was 15% or greater [4].

2.3 Chest radiography technique

The mobile chest radiography unit used in this study is a "FUJIFILM FDR nano" MOBILE X-RAY UNIT DR-XD 1000PX. It can generate a tube voltage of 40 to 100 kVp and a tube current of up to 35 mA. Upright patient's position and posteroanterior view were obtained.

2.4 Reference standard

All chest radiographs were analyzed by one thoracic subspecialty radiologist (more than 10 years of experience in thoracic imaging) and one in-training radiology fellow in comparison with the prior interpretation by the other radiologist. Both the thoracic radiologist and the in-training fellow were blinded to the results of the DLAD. All cases were assessed in consensus for negative findings, pneumonia, and other findings.

2.5 Statistical analysis

Statistical analysis was performed with STATA (version 14.0). The sample size was estimated to be 820, based on the expected sensitivity and specificity, which were 80% and 70% respectively [11]. One-sample comparison of proportion formula was performed with a 95% confidence interval, error (alpha) 0.05, and an assumption that prevalence of pneumonia was 30%. Descriptive statistics are reported as mean and standard deviation (SD), or median and interquartile range (IQR) for continuous data. Counts and percentages were performed for categorical data. Statistical differences in the demographics were evaluated using Student's t-test or Mann-Whitney U test for quantitative variables and Fisher's exact test for qualitative variables. The diagnostic performance in the detection of COVID-19 pneumonia was evaluated by constructing the receiver operating characteristic curve (ROC) and calculating the area under the curve (AUC) together with the sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV). The effects of age, sex, BMI, and the duration from the onset of symptoms to the time of chest radiography on COVID-19 pneumonia were assessed using multivariable logistic regression analysis. All statistical analyses were two-sided, and significance was assigned at P-value < 0.05.

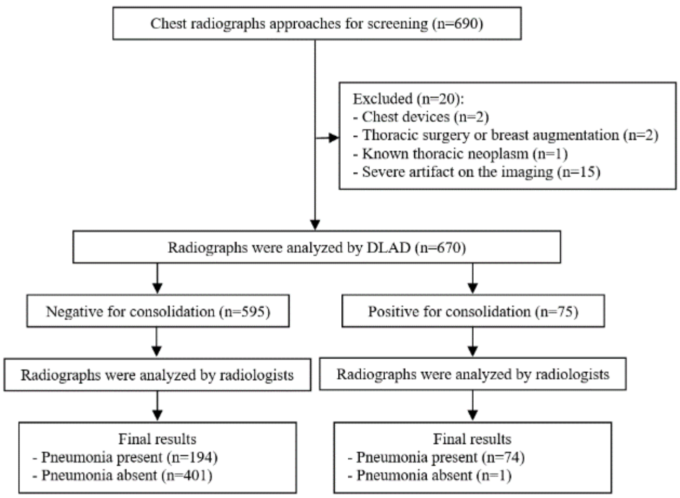

Fig. 1. Flow diagram of the study and final results.

3. Results and Discussion

3.1 Patient demographics

Between 1 May 2021 and 31 July 2021, we enrolled 690 patients for initial eligibility. Twenty patients were excluded due to indwelling chest devices (*n*=2), thoracic surgery or breast augmentation (*n*=2), known thoracic neoplasm (*n*=1), and severe artifacts on the image (*n*=15). The remaining 670 patients were included in the study and their chest radiographs were analyzed by the DLAD and the radiologists with a completion rate of 97% (Fig. 1). The included 670 patients consisted of 277 men (41%) and 393 women (59%) with a mean age of 38.01 years (SD 14.22). A mean BMI was 24.95 kg/m2 (SD 5.39). All patients were asymptomatic or mildly symptomatic according to the criteria for admission into the Thammasat University Field Hospital. Five hundred twenty-two (78%) patients presented with one or more of the symptoms of fever, cough, loss of taste or smell, sore throat, or mild dyspnea. One hundred and forty-eight (22%) asymptomatic patients were enrolled for the RT-PCR test due to a history of contact with a known case of COVID-19 infection or were admitted into the Thammasat University Field Hospital through the active screening networks in the outbreak areas. The prevalence of COVID-19 pneumonia in the study was 40% (268/670) (95%CI 36%-43.8%). In crude analysis, age, BMI, symptoms, and symptom onset to chest radiography were significantly different between the groups with and without pneumonia, while sex distribution was not significantly different (Table 1).

Table 1. Baseline demographic characteristics of patients with COVID-19 with and without pneumonia.

Variables	All patients	Patients with pneumonia	Patient without pneumonia	*P*-value
Number of patients	670	267 (39.85)	403 (60.15)	-
Gender				
Male, n (%)	277 (41.34)	108 (40.30)	169 (42.04)	-
Female, n (%)	393 (58.66)	160 (59.70)	233 (57.96)	0.689
Age, years, mean (SD)	38.01 (14.22)	43.55 (14.44)	34.31 (12.82)	<0.001*
BMI, kg/m^2, mean (SD)	24.95 (5.39)	26.19 (5.39)	24.13 (5.25)	<0.001*
Symptomatic patients				
Asymptomatic, n (%)	148 (22.09)	45 (16.79)	103 (25.62)	-
Symptomatic, n (%)	522 (77.91)	223 (83.21)	299 (74.38)	0.008*
The duration from the onset of symptoms to the time of chest radiography, days, median (IQR)	4 (2-7)	6 (2-6)	6 (3-9)	0.020*
Others radiologic findings				-
Pleural effusion, n	0	0	0	
Pneumothorax, n	0	0	0	
Calcification, n	1	0	1	
Nodule or mass, n	3	0	3	
Tuberculosis, n	7	2	5	

*Statistically significant at *P*-value < 0.05 determined by student's t-test, Mann-Whitney U test, and Fisher's exact test

3.2 Diagnostic performance of DLAD for pneumonia detection

The cross-tabulation of the results of the DLAD and radiologists are shown in Table 2. The radiologists analyzed those 268 patients that had pneumonia. The sensitivity and specificity of DLAD in the identification of pneumonia were 27.6% (95%CI 22.3%-33.4%) and 99.8% (95%CI 98.6%-100%), respectively. PPV and NPV were 98.7% (95%CI 92.8%-100%) and 67.4% (95%CI 63.5%-71.2%), respectively, with an AUC of 0.64 (95%CI 0.61-0.66). The ROC is shown in Fig. 2. Median and IQR of the abnormality score for pneumonia were 56% and 37%-65%

respectively. Examples of these radiographs are shown in Figs. 3-7.

Table 2. DLAD performance with a diagnosis of COVID-19 pneumonia.

	Radiologist		
DLAD	Positive	Negative	Total
Positive	74	1	75
Negative	194	401	595
Total	268	402	670

There were 194 false negative radiographs. Most cases with ground-glass opacity interpreted by the radiologists were not detected by the DLAD. This also led to underestimation of the distribution and severity of pneumonia by the DLAD. The only one false positive case was a 28-year-old woman whose radiograph showed reticulonodular opacity at the right upper lung zone which is likely considered to represent an active tuberculosis.

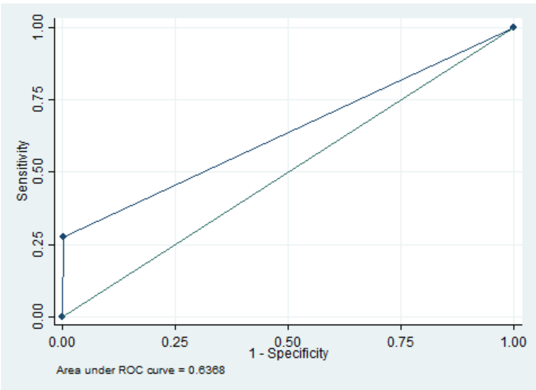

Fig. 2. ROC of DLAD for COVID-19 pneumonia detection.

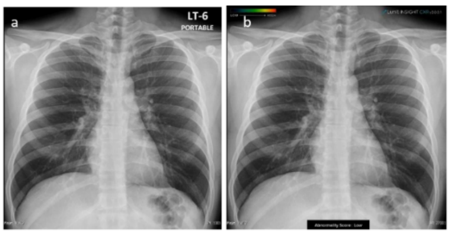

Fig. 3. An example of a true negative radiograph in a 36-year-old man. (a) Chest radiograph showed no detected abnormality in both lungs. (b) DLAD revealed a low probability of abnormality.

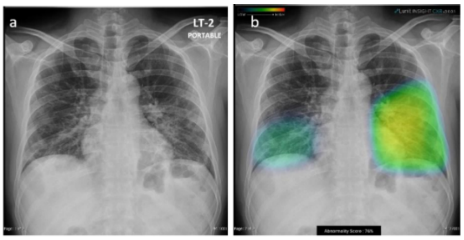

Fig. 4. An example of a true positive radiograph in a 40-year-old man. (a) Chest radiograph showed multifocal ground-glass opacities in both mid to lower lung zones, consistent with pneumonia. (b) DLAD interpreted as multiple consolidations with an abnormality score of 76%.

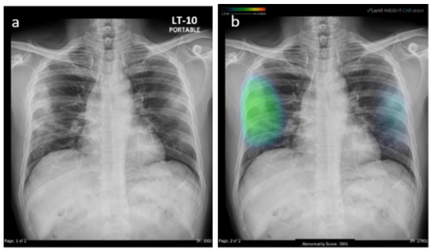

Fig. 5. An example of a true positive radiograph in a 55-year-old man. (a) Chest radiograph showed multifocal ground-glass opacities in the periphery of both mid to lower lung zones, consistent with pneumonia. (b) DLAD interpreted as multiple consolidations with an abnormality score of 56%.

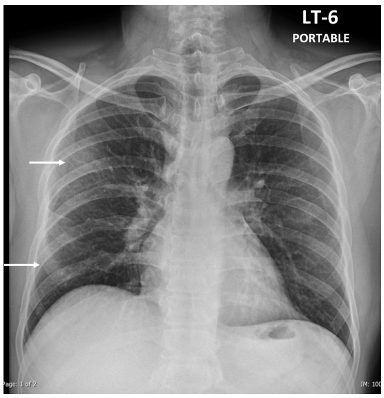

Fig. 6. An example of a false negative radiograph in a 48-year-old man. Chest radiograph showed subpleural ground-glass opacities in right mid to lower lung zones (arrows). DLAD classified as the low probability for abnormality.

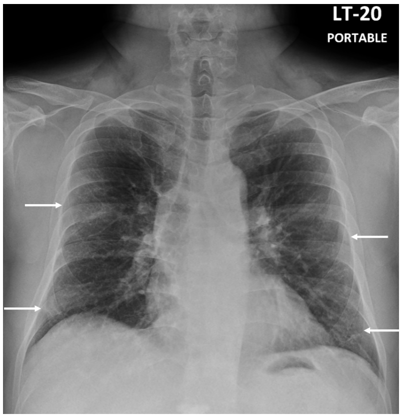

Fig. 7. An example of a false negative radiograph in a 42-year-old man. Chest radiograph showed multifocal subpleural ground-glass opacities in both mid to lower lung zones (arrows). DLAD classified as the low probability for abnormality.

3.3 Other imaging features

The other radiologic features on the radiologist report were calcification (*n*=1), nodule or mass (*n*=3), and pulmonary tuberculosis (*n*=7). The DLAD also reported 4 cases of pneumothorax and 1 case of pneumoperitoneum which were not found in the radiologists' analysis. However, the performance of the DLAD in the detection of these findings was not the purpose of this study. Examples of these radiographs are shown in Figs. 7-8.

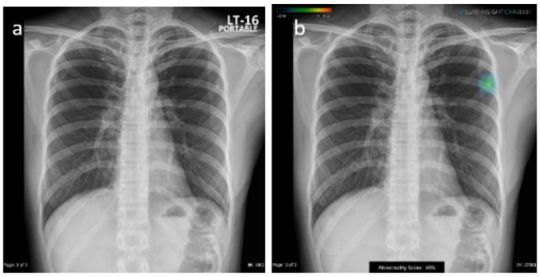

Fig. 8. (a) The chest radiograph of a 60-year-old woman with multiple masses and nodules in right lower lung zone. (b) The DLAD can also detect these lesions with an abnormality score of 98%.

Fig. 9. The DLAD falsely interpreted as pneumothorax at left upper lung zone with an abnormality score 33% but there was actually no pneumothorax.

3.4 Effects of demographics on COVID-19 pneumonia

Age, BMI, symptoms, and the duration from the onset of symptoms to the time of chest radiography were significantly different between the two groups in crude analysis. By multivariable logistic regression analysis, only the duration from the onset of symptoms to the time of chest radiography was a significant predictor of COVID-19 pneumonia (*P*-value=0.013), whereas age, BMI, and symptoms were not (Table 3).

Table 3. Multivariable regression analysis of variables predicting pneumonia.

Variables	Odds ratio (95%CI)	*P*-value
Age	1.02 (0.99-1.05)	0.236
BMI	1.05 (0.96-1.15)	0.259
Symptomatic patients (Reference: asymptomatic patient)	0.10 (0.01-1.24)	0.073
The duration from the onset of symptoms to the time of chest radiography	1.20 (1.04-1.39)	0.013*

* Statistically significant at *P*-value < 0.05 determined by multivariable logistic regression.

3.5 Discussion

The purpose of this study was to validate the diagnostic performance in COVID-19 pneumonia detection of the DLAD for mobile chest radiographs in patients with SARS-CoV-2 infection admitted into the Thammasat University Field Hospital. The study has demonstrated that the DLAD has excellent specificity and PPV (almost 100%) but low sensitivity (about 28%) and AUC (0.64).

The prevalence of pneumonia in COVID-19 patients in our study (40%) was slightly lower than the prior study by Inui et al [5] on the group of patients from the cruise ship “Diamond Princess”. Inui found that 54% of asymptomatic patients and 79% of symptomatic patients had lung opacities on CT. This is probably due to the patients’ demographics, the lower sensitivity of chest radiography as compared with CT scan, and possibly the virus present being a different variant.

Jang et al. [6] evaluated the diagnostic performance of deep-learning algorithms in patients who were admitted to five emergency departments and one community treatment center in Korea, and found the sensitivity, specificity, and AUC to be 95.6%, 88.7%, and 0.921 respectively.

Hwang et al. [4] found that radiologists with deep learning-based computer-aided detection assistance could identify COVID-19 pneumonia on chest radiographs with a sensitivity of 81.5% and a specificity of 72.3%.

The sensitivity and AUC of our study were lower than those of the previous studies because our study included patients that tended to have milder forms of pneumonia and lower density of air-space lesion on their chest radiographs. Our findings are concordant with the findings of Inui et al. [5] which indicate asymptomatic cases would more likely show ground-glass opacity than consolidation (83%) while symptomatic cases would more frequently show consolidation than ground-glass opacity (41%). Interpretation by radiologists in consensus was used to increase the accuracy of COVID-19 pneumonia diagnosis. However, in some cases, radiologists may overinterpret, leading to a decrease in the sensitivity of the DLAD. Even though DLAD cannot detect mild or faint ground glass opacity, it might not significantly affect clinical outcomes and management. In clinical practice, patients have other parameters for clinical decisions, such as symptoms, laboratory results, or serial chest radiographs, which contribute to a more comprehensive assessment of their condition.

Excellent specificity and PPV of the DLAD imply that if the DLAD suggests that a patient has COVID-19 pneumonia, the patient typically has the actual disease. These results may improve the quality of diagnosis, increase confidence in the findings, and reduce the time required for imaging interpretation, especially in healthcare settings with a shortage or lack of radiologists.

The systematic review by Wang et al. [7] on the application of artificial intelligence (AI) techniques in COVID-19 revealed that the AI achieved high performance in diagnosis for COVID-19. In this review, 46 studies related to AI-assisted diagnosis through chest images for COVID-19 showed total accuracy of 70.00-99.92%, sensitivity of 73.00-100.00%, specificity of 25-100.00%, and an AUC of 0.732 to 1.000. The detection of pneumonia in known COVID-19 patients was not included in this review.

In another comparison with the same deep-learning algorithm in different diseases, Hwang et al. [3] validated the algorithm’s performance in four major thoracic diseases including pulmonary malignant neoplasm, active tuberculosis, pneumonia, and pneumothorax. A median (range) AUC of 0.979 (0.973-1.000) for image-wise classification and 0.972 (0.923-0.985) for lesion-wise localization were reported by Hwang et al. [3], and these were higher than the same parameters in our study on the detection of pneumonia in COVID-19 patients.

Lee et al. [8] studied multiple risk factors of pneumonia development in mild COVID-19 including age, sex, BMI, comorbidities, symptoms, and laboratory results. This study found that risk factors for pneumonia were age 65 years or more (*P*-value = 0.0.27), cough (*P*-value = 0.003), dyspnea (*P*-value = 0.034) and diarrhea (*P*-value < 0.001). Similar to our study, Lee et al. found that sex and BMI were not significant risk factors. However, in our study, we did not include patients older than 70 years old, therefore the age was not a significant variable.

The duration from the onset of symptoms to the time of chest radiography was the only one significant variable predicting pneumonia in our study with an odds ratio of 1.2 (95%CI 1.04-1.39, *P*-value = 0.013). Shorter duration of onset of symptoms to the time of chest radiography was likely associated with decreased possibility of developing pneumonia, corresponding with temporal changes of the disease process.

3.6 Limitations

Firstly, chest radiographs with radiologists' interpretation were the reference standards in our study. CT is more sensitive than chest radiography in the detection of pneumonia, especially in the early stage of disease [2]. However, CT could not be used for all patients with COVID-19 in clinical practice due to lack of indications for CT scan, infectious control measures, availability, and cost. Chest radiographs in our study were analyzed by one thoracic subspecialty radiologist (more than 10 years of experience in thoracic imaging) and one fellowship-in-training radiologist in consensus and comparison with the prior interpretation by the other radiologist to increase the accuracy.

Secondly, we focused on the final diagnosis of COVID-19 pneumonia in chest radiographs which could alter patient management, but we did not collect the data of severity and distribution of the disease. Further study of the performance of the DLAD in severity evaluation may have prognostic benefits.

Thirdly, more clinical factors may have effects on the development of pneumonia such as patients' comorbidities, details and severity of symptoms, smoking, variants of virus, and history of vaccination. Due to the nature of retrospective studies, we could not collect all the above-mentioned expected predicting factors.

Fourthly, there may be technical factors affecting the DLAD's efficacy, such as X-ray generator settings or imaging quality, which were not included in this study.

These limitations will be taken into consideration in our future studies.

4. Conclusion

DLAD is a tool with excellent specificity and PPV for pneumonia detection but exhibits low sensitivity, especially in cases of ground-glass opacity on chest radiographs. A negative result on the DLAD may indicate either no pneumonia or a mild case, but a positive result on the DLAD should raise concerns about pneumonia. The duration from the onset of symptoms to the time of chest radiography is a significant predictor of pneumonia.

References

[1] Speets AM, van der Graaf Y, Hoes AW, Kalmijn S, Sachs AP, Rutten MJ, Gratama JW, Montauban van Swijndregt AD, Mali WP. Chest radiography in general practice: indications, diagnostic yield and consequences for patient management. Br J Gen Pract, 2006;56:574-8.

[2] Geoffrey DR, Christopher JR, Linda BH, Nicola S, Jeffrey PK, Suhail R, Neil WS, Annalisa V, Jae-Joon Y, Ian BK, Martin, Deverick JA, Christina K, Talissa A, Andrew B, Sujal RD, onathan G, Jin MG, Marc H, Yoshikazu I, Hans-Ulrich K, Fengming L, Peter JM, Mathias P, Martine RJ, Luca R, Cornelia MS-P, Noriyuki T, Athol UW, Ann NL. The Role of Chest Imaging in Patient Management during the COVID-19 Pandemic: A Multinational

Consensus Statement from the Fleischner Society. Radiology, 2020;296:172-80.

[3] Eui JH, Sunggyun P, Kwang-Nam J, Jung IK, So YC, Jong Hyuk L, Jin MG, Jaehong A, Jae-Joon Y, Julien GC, Gilbert RF, Chang MP. Development and Validation of a Deep Learning–Based Automated Detection Algorithm for Major Thoracic Diseases on Chest Radiographs. JAMA Netw Open, 2019;2(3):e191095.

[4] 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;21:1150-60.

[5] Inui S, Fujikawa A, Jitsu M, Kunishima N, Watanabe S, Suzuki Y, Umeda S, Uwabe Y. Chest CT Findings in Cases from the Cruise Ship Diamond Princess with Coronavirus Disease (COVID-19). Radiol Cardiothorac Imaging, 2020;17;2(2):e200110.

[6] Jang SB, Lee SH, Lee DE, Park SY, Kim JK, Cho JW, Cho J, Kim KB, Park B, Park J, Lim JK. Deep-learning algorithms for the interpretation of chest radiographs to aid in the triage of COVID-19 patients: A multicenter retrospective study. PLOS ONE, 2020;15(11):e0242759.

[7] Wang L, Zhang Y, Wang D, Tong X, Liu T, Zhang S, Huang J, Zhang L, Chen L, Fan H, Clarke M. Artificial Intelligence for COVID-19: A Systematic Review. Front Med (Lausanne), 2021;30;8:704256.

[8] Lee HW, Yoon SY, Lee JK, Park TY, Kim DK, Chung HS, Heo EY. Clinical implication and risk factor of pneumonia development in mild coronavirus disease 2019 patients. Korean J Intern Med, 2021;36:1-10.