



Applications of artificial intelligence in the thorax: a narrative review focusing on thoracic radiology

Yisak Kim^{1,2^}, Ji Yoon Park³, Eui Jin Hwang^{3^}, Sang Min Lee⁴, Chang Min Park^{1,2,3,5^}

¹Interdisciplinary Program in Bioengineering, Graduate School, Seoul National University, Seoul, Korea; ²Integrated Major in Innovative Medical Science, Seoul National University Graduate School, Seoul, Korea; ³Department of Radiology, Seoul National University College of Medicine, Seoul, Korea; ⁴Departments of Radiology and Research Institute of Radiology, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Korea; ⁵Institute of Medical and Biological Engineering, Medical Research Center, Seoul National University, Seoul, Korea

Contributions: (I) Conception and design: Y Kim, CM Park; (II) Administrative support: CM Park; (III) Provision of study materials or patients: EJ Hwang, SM Lee; (IV) Collection and assembly of data: All authors; (V) Data analysis and interpretation: Y Kim, JY Park; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Chang Min Park, MD, PhD. Department of Radiology and Institute of Radiation Medicine, Seoul National University College of Medicine, 101 Daehak-ro, Jongno-gu, Seoul 03080, Korea. Email: cmpark.morphius@gmail.com.

Objective: This review will focus on how AI—and, specifically, deep learning—can be applied to complement aspects of the current healthcare system. We describe how AI-based tools can augment existing clinical workflows by discussing the applications of AI to worklist prioritization and patient triage, the performance-boosting effects of AI as a second reader, and the use of AI to facilitate complex quantifications. We also introduce prominent examples of recent AI applications, such as tuberculosis screening in resource-constrained environments, the detection of lung cancer with screening CT, and the diagnosis of COVID-19. We also provide examples of prognostic predictions and new discoveries beyond existing clinical practices.

Background: Artificial intelligence (AI) has shown promising performance for thoracic diseases, particularly in the field of thoracic radiology. However, it has not yet been established how AI-based image analysis systems can help physicians in clinical practice.

Methods: This review included peer-reviewed research articles on AI in the thorax published in English between 2015 and 2021.

Conclusions: With advances in technology and appropriate preparation of physicians, AI could address various clinical problems that have not been solved due to a lack of clinical resources or technological limitations.

Keywords: Artificial intelligence (AI); deep learning (DL); computer aided diagnosis (CAD); thoracic radiology; pulmonary medicine

Submitted Aug 15, 2021. Accepted for publication Dec 14, 2021.

doi: 10.21037/jtd-21-1342

View this article at: <https://dx.doi.org/10.21037/jtd-21-1342>

Introduction

Artificial intelligence (AI) techniques have shown promising performance in medicine, particularly in the field of medical image analysis. Convolutional neural network (CNN)-

based deep learning (DL) models have shown performance equal to or even surpassing that of experts in various tasks, including the detection of retinal pathologies in fundus photographs (1-4), interpretation of echocardiography (5-7) and screening mammography (8), and the diagnosis of major

[^] ORCID: Yisak Kim, 0000-0002-5420-1436; Eui Jin Hwang, 0000-0002-3697-5542; Chang Min Park, 0000-0003-1884-3738.

thoracic diseases on chest radiographs (CXR) and computed tomography (CT) (9-27).

Radiology in respiratory medicine is a particularly important field for which a variety of AI applications are actively being developed. Major thoracic diseases such as lung cancer and tuberculosis are among the leading causes of death worldwide (28,29), and numerous radiologic studies have been performed to diagnose them. For example, CXR, which is often the first imaging study acquired to diagnose pathologies in the thorax, remains the most commonly performed radiologic exam worldwide, with an average of 238 CXRs acquired per 1,000 population annually (30). For this reason, even though DL research in medicine is notoriously data-hungry, DL in the thoracic radiology field, which has relatively large imaging databases, has been actively researched.

Recently, Nam *et al.* showed that a DL algorithm detecting 10 common abnormalities on CXR could improve radiologists' performance and shorten the reporting time for critical and urgent cases (18). Similarly, Seah *et al.* showed that a DL algorithm significantly improved the classification accuracy of radiologists for 102 clinical findings (31). In 2021, more than 13 US Food and Drug Administration–cleared AI algorithms developed for pulmonary diseases are available (32), and they are expected to be implemented sooner or later in daily clinical practice.

This review will focus on how AI (specifically, DL) can be applied to complement aspects of the current healthcare system. We included peer-reviewed research articles on AI in the thorax published in English between Jan 2015 and July 2021. A PubMed literature search performed on July 16, 2021, using the search terms “(artificial intelligence OR machine learning OR deep learning) AND (thorax OR pulmonary OR respiratory OR chest OR lung) AND medicine”. Under these search terms, more than 3,600 papers were searched, so given the narrative nature of this review, articles were carefully selected by reviewing their title and abstracts to provide a general understanding of this topic (Table 1).

The rest of the paper is organized as follows. In “Application schemes of AI tools in clinical practices” section, we describe how AI-based tools can augment existing clinical workflows by discussing the applications of AI to worklist prioritization and patient triage, the performance-boosting effects of AI as a second reader, and the use of AI to facilitate complex quantifications. We also introduce prominent examples of recent AI applications, such as tuberculosis screening in resource-

constrained environments, the detection of lung cancer with screening CT, and the diagnosis of coronavirus disease 2019 (COVID-19) in “Potential examples of AI-assisted clinical practice for thoracic diseases practices” section. We also provide examples of prognostic predictions and new discoveries beyond existing clinical practices in the “Prognostic prediction and new discoveries” section. Lastly, we close our review with a discussion of challenges and future directions of AI applications in thoracic radiology.

We present the following article in accordance with the Narrative Review reporting checklist (available at <https://dx.doi.org/10.21037/jtd-21-1342>).

Application schemes of AI tools in clinical practices

It is now well known that AI shows expert-level performance in interpreting medical images, but how AI-based tools can help physicians in clinical practice has not been yet established. One of the classic ways of integrating AI-based tools into the existing clinical workflows is the triage scenario, in which an AI system makes a provisional analysis and prioritizes the worklist in terms of the urgency of detected findings (11). It has been mainly investigated in the field of emergency medicine, in which the timely diagnosis and management of acute diseases are critical. Another major use of AI in clinical workflows is the add-on scenario (11). In this scenario, physicians check the results of the AI system during or after their interpretation to improve their diagnostic performance. The results of the AI system could include the probability value of a certain radiologic study being abnormal and localization of a specific disease. Alternatively, the results could be quantification values for further quantitative analysis of pathologies in imaging studies.

AI-based triage and worklist prioritization

Worklist prioritization is an important application of AI in thoracic radiology. It is clinically relevant, especially in the ED, where the timely diagnosis and management of acute diseases can often be critical. In the United States, there were over 130 million total visits to EDs in 2011, accounting for approximately 15% of all hospital visits (45-47). As pneumonia and respiratory symptoms have become one of the most common reasons for ED visits, radiologic studies—particularly CXR, a primary examination tool for evaluations in the ED—have shown a significant increase in use over the past two decades (48). Since the number of physicians,

Table 1 Cohort size, validation, algorithm type and reported performance for selected studies in thoracic radiology

First author (ref.)	Journal (year)	Cohort sizes	Source of data	Algorithm type	Task	Performance
Nam (18)	<i>Eur Respir J</i> (2021)	146,717 CXRs from 108,053 patients	Data collected from Seoul National University Hospital	ResNet34-based	Detect 10 common abnormalities on CXR	AUROC 0.895–1.00 in the CT-confirmed external dataset and 0.913–0.997 in the PadChest
Seah (31)	<i>Lancet Digit Health</i> (2021)	821,681 CXRs from 284,649 patients	MIMIC, I-MED, ChestX-ray14, CheXpert, and PadChest	EfficientNet-based for classification and U-Net-based for segmentation	CXR interpretation across 127 clinical findings	AUROC 0.954–0.959 in the MIMIC and I-MED
Huang (33)	<i>NPJ Digit Med</i> (2020)	1,797 CTPA studies from 1,773 patients	CTPA dataset collected from a single institution	3D CNN PENet	Detect PE on volumetric CTPA scans	AUROC of 0.82–0.87 on the hold out internal testset and 0.81–0.88 on external dataset
Hata (34)	<i>Eur Radiol</i> (2021)	170 non-contrast-enhanced CT from 170 patients	Data collected from single institution	Xception-based	Detect AD on non-contrast-enhanced CT	Accuracy, sensitivity, and specificity of 90.0%, 91.8%, and 88.2%
Hwang (16)	<i>Radiology</i> (2019)	89,834 CXRs for train and 1,135 patients for validation	Data collected from Seoul National University Hospital	DenseNet-based	Detect four major thoracic diseases on CXRs	AUROC of 0.93–0.96 for validation dataset
Hasenstab (35)	<i>Radiol Cardiothorac Imaging</i> (2021)	CT from 8,951 patients	Data collected from the COPD Genetic Epidemiology study	Deep CNN	Stage the severity of COPD through quantification of CT	Stages correlated with the GOLD criteria, with AUROC of 0.86–0.96
Chassagnon (36)	<i>Radiol Artif Intell</i> (2020)	CT from 208 patients	Data collected from single institution	SegNet autoencoder-based AtlasNet	Assessment of the extent of systemic sclerosis related ILD	Dice similarity coefficients of 0.74–0.75 for ILD contours
Hwang (20)	<i>Clin Infect Dis</i> (2019)	60,989 CXRs from 50,593 patients	Data collected from Seoul National University Hospital and 6 external multicenter or validation	27-layer deep CNN	Detect active pulmonary tuberculosis on CXRs	AUROC of 0.977–1.000 for classification and AUAFROC of 0.973–1.000 for localization in external dataset
Ciampi (37)	<i>Sci Rep</i> (2017)	1,805 nodules from 943 patients	Data from the Multicentric Italian Lung Detection trial	Deep CNN	Classifying lung nodules into 6 classes	Average accuracy of 72.9%
Ardila (38)	<i>Nat Med</i> (2019)	42,290 CT from 14,851 patients	Data from the National Lung Cancer Screening Trial	Mask-RCNN, RetinaNet, Inception V1 and 3D Inception	Predict the risk of lung cancer based on CT	AUROC of 0.94

Table 1 (continued)

Table 1 (continued)

First author (ref.)	Journal (year)	Cohort sizes	Source of data	Algorithm type	Task	Performance
Harmon (39)	<i>Nat Commun</i> (2020)	CT from 1,280 patients for training and 1,337 patients for validation	Data from four international centers	AH-Net and DenseNet 121-based	Detect COVID-19 pneumonia on CT	Accuracy, sensitivity, and specificity of 90.8%, 84%, and 93%
González (40)	<i>Am J Respir Crit Care Med</i> (2018)	CT from 7,983 COPDGene participants and 1,672 ECLIPSE participants	Data from COPDGene and ECLIPSE	Deep CNN	Acute respiratory disease event and mortality prediction on CT	Acute respiratory disease event prediction (C-index, 0.64 and 0.55 for internal and external validation) and mortality prediction (C-index, 0.72 and 0.60)
Hosny (41)	<i>PLoS Med</i> (2018)	CT from 1,194 patients	Data from 7 independent datasets across 5 institutions	3D CNN	2-year mortality prediction of NSCLC patients	AUROC of 0.70 and 0.71 for 2-year mortality after the start of radiotherapy and after surgery
Lu (42)	<i>JAMA Netw Open</i> (2019)	CXRs from 57,813 patients	Data from Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial and the National Lung Screening Trial	CXR-risk CNN	12-year mortality prediction from a CXR	The very high CXR-risk group had mortality of 53.0% (PLCO) and 33.9% (NLST)
Chao (43)	<i>Nat Commun</i> (2021)	CT from 10,730 patients	Data from the National Lung Screening Trial and Massachusetts General Hospital	Tri2D-Net-based	Predict cardiovascular mortality with low-dose-CT	AUROC of 0.734–0.801
Raghu (44)	<i>JACC Cardiovasc Imaging</i> (2021)	CXRs from 116,035 individuals	Data from Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial and the National Lung Screening Trial	Deep CNN	Estimate biological age from a CXR to predict longevity	CXR-Age carried a higher risk of all-cause mortality than a 5-year increase in chronological age

CXR, chest radiograph; CNN, convolutional neural network; COPD, chronic obstructive pulmonary disease; CTPA, computed tomography pulmonary angiography; NSCLC, non-small-cell lung cancer; ILD, interstitial lung disease; PE, pulmonary embolism; AD, aortic dissection.

including radiologists, is not sufficient compared to the increasing number of imaging tests, provisional analysis and prioritization by AI systems can directly improve the clinical outcomes of patients whose timely diagnosis is critical. Nam *et al.* (18) performed simulated reading tests for CXRs from ED patients with and without AI assistance, and they found that radiologists detected significantly more critical (29.2% to 70.8%) and urgent (78.2% to 82.7%) abnormalities when aided by the AI system; furthermore, AI assistance shortened the time-to-report for CXRs of critical and urgent cases (from 3,371.0 to 640.5 s and from 2,127.1 to 1,840.3 s, respectively). *Figure 1* shows an example of utilizing AI tools for worklist prioritization.

Prioritization of CT images is another important topic in the ED. For example, the usage of computed tomography pulmonary angiography (CTPA) in the ED to diagnose pulmonary embolism (PE) has increased 27-fold over the past two decades (49,50). A triage tool to automatically identify clinically important PEs and prioritize CT images of PE patients can improve care pathways via more efficient diagnoses. Huang *et al.* (33) developed a DL algorithm to automatically detect PE on volumetric CTPA scans as an end-to-end solution. Without requiring computationally intensive and time-consuming preprocessing, it achieved AUCs of 0.84 and 0.85 on detecting PE in the internal test set and external dataset, respectively.

Aortic dissection (AD) is another common emergency that is often fatal. Contrast-enhanced CT is the most commonly used diagnostic modality for AD (51), but detection and prioritization of acute AD on non-contrast-enhanced CT is also useful in the ED. Hata *et al.* (34) developed a DL algorithm to detect AD on non-contrast-enhanced CT. The developed DL algorithm provided accuracy, sensitivity, and specificity of 90.0%, 91.8%, and 88.2%, respectively, with a cutoff value of 0.400. For the radiologists, the median accuracy, sensitivity, and specificity were 88.8%, 90.6%, and 94.1%, respectively. There was no significant difference in performance between the DL algorithm and the average performance of the radiologists, demonstrating the potential of the DL algorithm for provisional analysis and prioritization of CT images in the ED.

AI as a second reader

Many recent studies have compared radiologists' performance in image interpretation with and without AI assistance. For identifying abnormalities on CXRs, such as active tuberculosis (20), malignant nodules (21), or major

thoracic diseases (19), AI assistance led to meaningful improvements in physician readers' performance. In this scheme, the AI system provides probability values of specific diseases with or without localization information, and physicians check these results during or after their image interpretation. An example of this type of clinical workflow is presented in *Figure 2*.

Nam *et al.* (18) developed a DL algorithm detecting 10 common abnormalities (pneumothorax, mediastinal widening, pneumoperitoneum, nodule/mass, consolidation, pleural effusion, linear atelectasis, fibrosis, calcification, and cardiomegaly) on CXRs, as well as providing a visual localization of the abnormality. Similarly, Seah *et al.* (31) conducted a study in which 20 radiologists reviewed CXRs across 127 clinical findings with and without the assistance of a DL algorithm, and found that radiologists assisted by the DL algorithm showed much better reading performances, with higher areas under the curve (AUCs) when assisted by the DL algorithm (AUC, 0.808; 95% CI, 0.763–0.839) than when not assisted (AUC, 0.713; 95% CI, 0.645–0.785). The DL algorithm significantly improved the classification accuracy of radiologists for 102 (80%) of 127 clinical findings and was statistically non-inferior for 19 (15%) findings; furthermore, no findings showed a decrease in accuracy when radiologists used the DL algorithm.

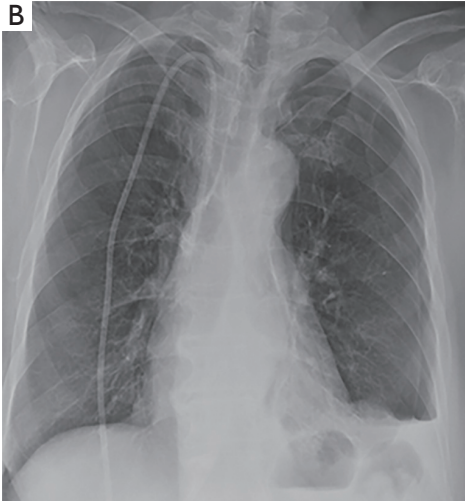
The added value of AI assistance is particularly prominent in specific situations such as emergencies. CXR is a simple and widely accessible imaging modality; however, its interpretation is not easy and often requires a high quality of expertise and experience. Many studies have found substantial discordances in CXR interpretation in the emergency department (ED), ranging from 0.3% to 17% (52–54). This kind of misinterpretation and discordant interpretation of critical cases could directly influence patients' clinical courses and outcomes. Furthermore, physicians in the ED often have limited time or opportunity to reach an on-call radiologist for consultations (55). Hwang *et al.* (16) investigated whether the application of a commercially available DL algorithm could enhance clinicians' reading performance for clinically relevant abnormalities on CXRs in the ED setting. Assistance from the DL algorithm improved the sensitivity of radiology residents' interpretation from 65.6% to 73.4%. Later, in 2020, Kim *et al.* (12) also reported that with the support of a DL algorithm, physicians' diagnostic performance for pneumonia improved (sensitivity: 53.2% to 82.2%; specificity: 88.7% to 98.1%).

It is commonly understood that AI assistance brings synergistic effects because it finds missed findings or

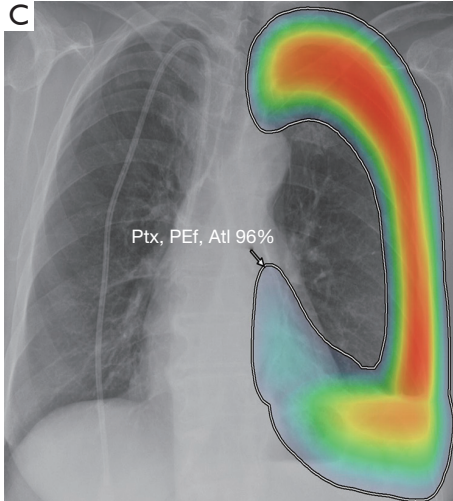
A

	Age	Sex	Protocol Name	Abnormal Finding	AI_Score	Urgency Finding	Urgency Score	Study DateTime
▶ 1	057Y	M	Chest PA	pneumothorax-29	57	pneumothorax-29	57	2021-08-10 11:02:40
▶ 2	059Y	M	Chest PA	pneumothorax-29	94	pneumothorax-29	94	2021-08-10 10:25:13
▶ 3	069Y	F	Chest PA	pleural_effusion-3	92	pneumothorax-29	59	2021-08-10 09:56:24
▶ 4	084Y	M	Chest PA	consolidation-7	80	pneumothorax-29	33	2021-08-10 09:50:06
▶ 5	063Y	M	Chest PA	pleural_effusion-3	92	pneumothorax-29	33	2021-08-10 09:16:02
▶ 6	061Y	F	Chest PA	pleural_effusion-3	98	pneumothorax-29	24	2021-08-10 06:52:51
▶ 7	056Y	F	Chest PA	pneumothorax-29	63	pneumothorax-29	63	2021-08-10 06:11:15
▶ 8	071Y	M	Chest PA	consolidation-7	43	pneumothorax-29	24	2021-08-10 06:06:49
▶ 9	054Y	M	Chest PA	pleural_effusion-3	98	pneumothorax-29	21	2021-08-10 06:04:23
▶ 10	057Y	M	Chest PA	pneumothorax-29	33	pneumothorax-29	33	2021-08-10 06:03:21
▶ 11	068Y	M	Chest PA	consolidation-7	83	pneumothorax-29	75	2021-08-10 06:02:11
▶ 12	082Y	M	Chest PA	consolidation-7	86	pneumothorax-29	35	2021-08-10 05:59:35
▶ 13	075Y	M	Chest PA	consolidation-7	90	pneumothorax-29	28	2021-08-10 05:57:33
▶ 14	079Y	M	Chest PA	pleural_effusion-3	88	pneumothorax-29	86	2021-08-10 05:56:49
▶ 15	072Y	M	Chest PA	pleural_effusion-3	89	pneumothorax-29	85	2021-08-10 05:56:08
▶ 16	043Y	F	Chest PA	pneumothorax-29	18	pneumothorax-29	18	2021-08-10 05:51:54
▶ 17	059Y	M	Chest PA	pneumothorax-29	92	pneumothorax-29	92	2021-08-10 05:50:06
▶ 18	073Y	M	Chest PA	pneumothorax-29	91	pneumothorax-29	91	2021-08-10 05:45:42
▶ 19	084Y	M	Chest PA	pleural_effusion-3	86	pneumothorax-29	29	2021-08-10 05:43:07
▶ 20	072Y	M	Chest PA	pneumothorax-29	92	pneumothorax-29	92	2021-08-10 05:39:20
▶ 21	055Y	F	Chest PA	pneumothorax-29	65	pneumothorax-29	65	2021-08-10 05:38:29
▶ 22	055Y	M	Chest PA	consolidation-7	87	pneumothorax-29	25	2021-08-10 05:37:50
▶ 23	060Y	M	Chest PA	pneumothorax-29	62	pneumothorax-29	62	2021-08-10 05:35:01
▶ 24	073Y	M	Chest PA	calcification-5	49	pneumothorax-29	32	2021-08-10 05:32:38
▶ 25	078Y	M	Chest PA	pneumothorax-29	97	pneumothorax-29	97	2021-08-10 05:31:46
▶ 26	056Y	F	Chest PA	pneumothorax-29	97	pneumothorax-29	97	2021-08-09 20:30:47
▶ 27	071Y	M	Chest PA	pneumothorax-29	23	pneumothorax-29	23	2021-08-09 17:57:40
▶ 28	072Y	M	Chest PA	pleural_effusion-3	93	pneumothorax-29	80	2021-08-09 16:36:00
▶ 29	069Y	F	Chest PA	pneumothorax-29	81	pneumothorax-29	81	2021-08-09 16:32:46
▶ 30	078Y	M	Chest PA	pneumothorax-29	98	pneumothorax-29	98	2021-08-09 16:27:53
▶ 31	061Y	F	Chest PA	pleural_effusion-3	85	pneumothorax-29	58	2021-08-09 16:09:28

B



C



Ptx, PEf, Atl 96%

Figure 1 Implementation of AI CAD into a PACS system for prioritization of chest radiographs. An AI-integrated PACS system can display the results of analysis by AI CAD on the exam list (A). It can provide not only the presence of any abnormal finding, but also the presence of urgent findings requiring timely interpretation (e.g., pneumothorax), along with the corresponding probability scores. An interpreting radiologist can sort chest radiographs by the presence and type of urgent findings or corresponding probability scores to interpret chest radiographs with urgent findings first. A chest radiograph of a 73-year-old female patient shows left hydro-pneumothorax (B). The AI CAD system identified the pneumothorax with a probability score of 96% (C). CAD, computer aided diagnosis; PACS, picture archiving and communication system.

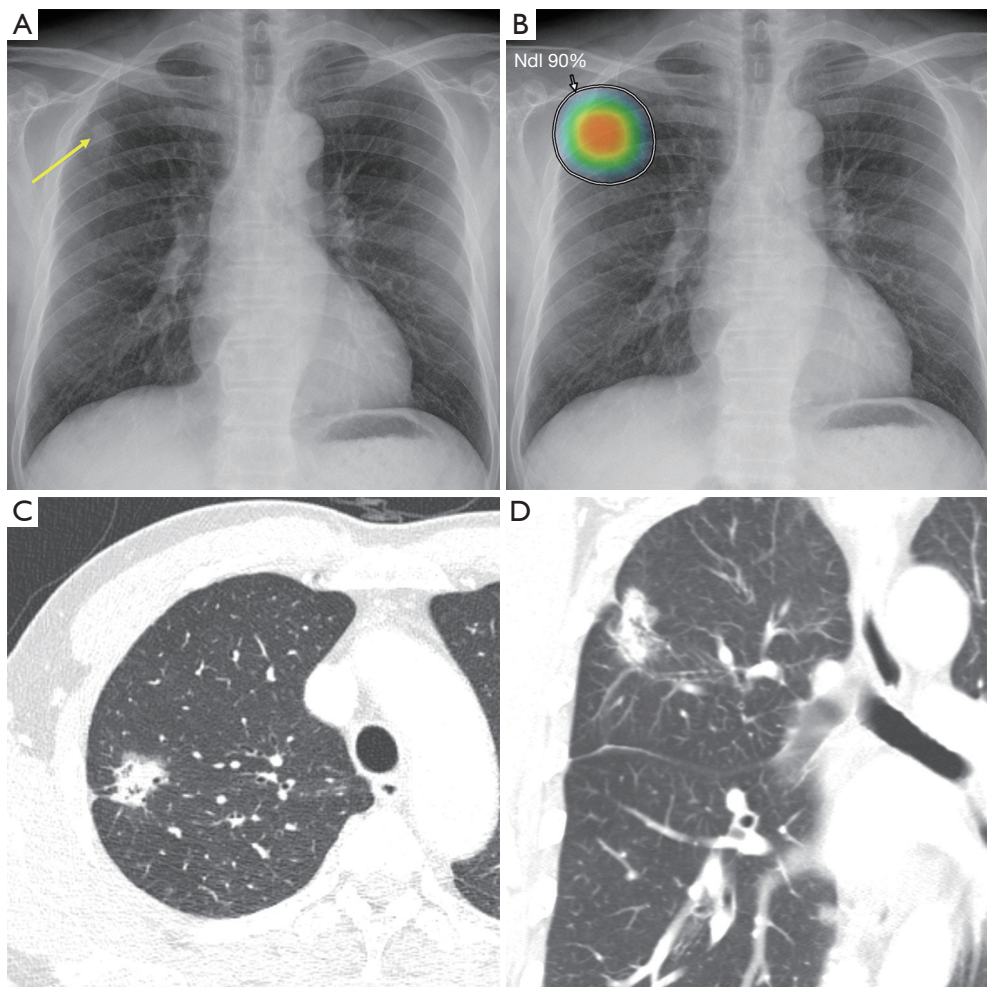


Figure 2 Identification of a lung nodule on chest radiograph using an AI CAD system A chest radiograph of a 71-year-old male patient shows a nodular opacity in the right upper lung field (A, arrow). The AI CAD system identified the nodule with a probability score of 90% (B). Chest CT of the patient shows an irregular nodule with air-bronchogram and a ground-glass opacity component in the right upper lobe of the lung (C,D). The nodule was proven to be lung cancer after surgery. CAD, computer aided diagnosis.

possible mistakes, like a second reader in double-reading. Double-reading is generally considered to be of added value in diagnostic radiology (56), and high-accuracy AI can act as a competent second reader to reduce perceptual errors (57). However, the interaction between AI systems and physicians is still poorly understood (58), and further research is needed to maximize the synergistic effects of AI.

Automatic quantification for complex quantitative analysis

AI applications in the add-on scenario can provide quantification results from medical images. Many studies have been conducted to find quantitative biomarkers from chest CT for lung cancer, chronic obstructive pulmonary

disease (COPD), and interstitial lung disease (ILD) (59-62). However, manual quantification is extremely time-consuming and practically impossible in routine clinical practice. In this context, AI-based quantification has been actively conducted, and several recent studies have shown that AI could improve the quantitative analysis in a highly accurate and time-efficient manner through automatic segmentation of lung parenchyma (63,64), pulmonary lobes (65), airways (66), and pulmonary pathologies (67,68). Hasenstab *et al.* (35) developed a DL algorithm to stage the severity of COPD through quantification of emphysema and air trapping on chest CT images. They proposed five CT-based COPD stages based on the percentage of emphysema and total lung involvement. The proposed stages correlated with the

predicted spirometry-based Global Initiative for Chronic Obstructive Lung Disease (GOLD) criteria, with AUCs of 0.86–0.96, and predicted disease progression (odds ratio, 1.50–2.67) and mortality (hazard ratio, 2.23; $P < 0.001$) both with and without the GOLD criteria. Similarly, Chassagnon *et al.* (36) developed a multicomponent deep neural network (AtlasNet), a DL algorithm for the automatic assessment of the extent of systemic sclerosis-related ILD on chest CT images. AtlasNet performed similarly to radiologists for disease-extent contouring, which is correlated with pulmonary function, to assess CT images from patients with systemic sclerosis-related ILD. The median dice similarity coefficients (DSCs) between the readers and the deep learning ILD contours ranged from 0.74 to 0.75, whereas the median DSCs between the contours from radiologists ranged from 0.68 to 0.71. *Figure 3* and *Figure 4* present examples of automatic quantification of COPD and ILD on CT, respectively.

Radiomics feature is another important quantitative biomarker that has recently emerged as a new field of radiologic research (69). There are studies that these measures can be strong indicators for lung cancer prognosis and phenotyping (70,71), but at the same time, challenging components of radiomics such as accurate segmentation and reproducibility over various devices institutions (69,71,72). DL is considered a promising method to solve these problems, as it has shown excellent performance in segmentation in chest CTs that can automate time-consuming manual segmentation, and ability to generates various image styles while maintaining the content that can improves radiomics reproducibility (73).

Potential examples of AI-assisted clinical practice for thoracic diseases

This section describes some prominent applications of AI-assisted systems for major thoracic diseases. We cover tuberculosis screening in resource-constrained environments, lung cancer detection on chest CT, and the diagnosis of COVID-19 through imaging studies.

Tuberculosis screening in resource-constrained environments

Although displaced by COVID-19 in 2020, tuberculosis was the leading cause of death among infectious diseases until 2019 (29). Because of the high proportion of undetected patients and the potential to reduce mortality through early

detection and treatment, the World Health Organization (WHO) has recommended systematic screening for tuberculosis for people at risk since 2013 (74). CXR is an effective screening tool for both children and adults due to its reasonably high sensitivity and specificity for tuberculosis detection (75).

However, in low- and middle-income countries (LMICs), CXR often shows lower sensitivity and specificity for tuberculosis detection than expected, which is related to the lack of well-trained radiologists. In fact, the sensitivity and specificity of tuberculosis diagnosis through CXR in a Nepalese center were 78% and 51%, respectively, and in Yogyakarta, Indonesia, the sensitivity and specificity were 88.6% and 82.9% (76,77). Furthermore, van't Hoog *et al.* compared the sensitivity and specificity of each tuberculosis screening method in seven countries including Kenya, Cambodia, and Vietnam. Although CXR outperformed symptom-based screening, there was substantial variation across countries (78).

Various attempts have been made to overcome this limitation, and computer-aided diagnosis (CAD) has emerged as a potential solution for tuberculosis screening. Because tuberculosis presents heterogeneous radiologic findings, early AI models with human-derived features did not show satisfactory performance and thus were not applied in practice for screening. However, in recent years, the application of DL has led to remarkable improvements in tuberculosis screening models. In 2016, Hwang *et al.* (79) developed a tuberculosis screening model by applying a deep CNN, which was in the spotlight in the image processing field at the time. They trained the model with 10,848 CXR images and tested it with datasets from Korea, the United States, and China to obtain AUCs of 0.964, 0.88, and 0.93, respectively. In 2019, Hwang *et al.* (20) developed a 27-layer deep CNN model and validated it with six external multicenter, multinational datasets. The created model showed sensitivity of 94.3–100% and specificity of 91.1–100%, with significantly higher performance in both classification and localization than a group of physicians consisting of non-radiology physicians, board-certified radiologists, and thoracic radiologists. *Figure 5* shows an example of identifying tuberculosis using a DL-based AI solution in clinical practice.

Driven by these advances in deep learning technology, the updated WHO guideline in 2020 recommended the use of CAD for tuberculosis screening for individuals aged 15 years and older in populations in which tuberculosis screening is recommended (75). The performance of



Figure 3 Fully automated quantification of non-enhanced chest CT using AI software in a 67-year-old man with chronic obstructive pulmonary disease (FEV1/FVC =33%, FEV1 =37%). (A) Axial, sagittal, coronal, and volume rendering images of fully automated lung and lobe segmentation results using an AI engine (Aview, version 1.1.39.6; Coreline Soft, Seoul, South Korea). (B) Coronal image with a LAA (under -950 HU) mask (red box) and a results table (red box) of the quantification analysis, with results such as volume, the LAA under -950 HU, mean lung density, and percentile index based on the lung and lobe segmentation. (C) Volume-rendering image (red box) and results table (yellow box) of a segmented airway with quantification results, including bronchus level, wall thickness, wall area, wall area percent, lumen diameter, lumen area, and tapering ratio. FEV1/FVC, forced expiratory volume in one second/forced vital capacity; LAA, low attenuation area.



Figure 4 Fully automated quantification of non-enhanced chest CT using AI software in 66-year-old man with usual interstitial pneumonia (FEV1/FVC =74%, FVC =53%, FEV1 =56%). (A) Axial, sagittal, coronal, and volume-rendering images of fully automated lung and lobe segmentation results using an AI engine (Aview, version 1.1.39.6; Coreline Soft, Seoul, South Korea). (B) Axial images with/without a lung texture segmentation mask [red = honeycombing (H), orange = reticular opacity (R), cyan = ground glass opacity (G), blue = consolidation (C), yellow = emphysema (E)]. (C) Pie chart (red box) and results table (yellow box) with quantification of texture analysis, based on the lung and lobe segmentation. FEV1/FVC, forced expiratory volume in one second/forced vital capacity.

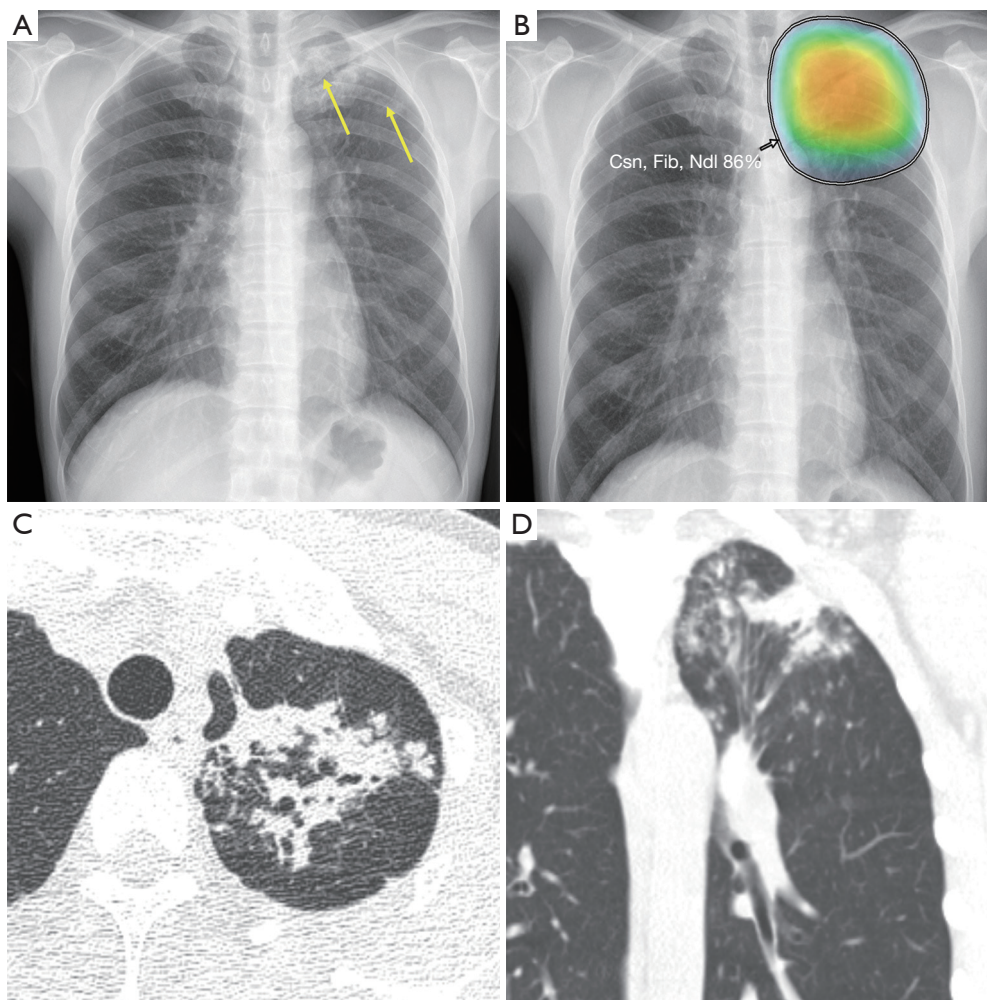


Figure 5 Identification of a chest radiograph from a patient with active pulmonary tuberculosis using an AI CAD system. (A) chest radiograph of a 52-year-old male patient with a cough shows clustered consolidation and nodules at the left lung apex (A, arrows). The AI CAD system identified the lesion with a probability score of 86% (B). Chest CT of the patient shows irregular consolidation and micronodules with bronchiectasis in the left upper lobe of the lung (C,D). The patient was diagnosed with active pulmonary tuberculosis by sputum acid-fast bacilli culture. CAD, computer aided diagnosis.

commercialized CAD systems was found to be non-inferior to that of physicians when applied in various regions, including LMICs. Qin *et al.* (80) tested three deep learning systems with the Xpert MTB/RIF assay-proven dataset from Nepal and Cameroon and obtained AUCs of 0.92, 0.94, and 0.94, respectively. When the sensitivity was matched with that of the radiologists, the specificity of two out of three systems was significantly higher than that of the radiologists. Khan *et al.* (81) applied commercial DL algorithms in Pakistan and showed that the sensitivity and specificity satisfied the WHO guidelines of 90% and 70%, respectively. CAD in LMICs can perform tuberculosis

screening, thereby improving health equity and accessibility and reducing mortality due to tuberculosis.

Lung cancer screening programs

In 2020, 2.21 million people were newly diagnosed with lung cancer, making it the second most diagnosed cancer after breast cancer. The number of cancer deaths due to lung cancer was 1.79 million, the highest among all cancers (28). The US National Lung Screening Trial (NLST) research team found that screening for lung cancer with low-dose CT (LDCT) in high-risk populations could reduce mortality

from lung cancer by 20% (82). The US Preventive Services Task Force currently recommends lung cancer screening with LDCT for adults aged 50 to 80 years with a smoking history of 20 pack-years or more, who are currently smoking or who have quit smoking within the past 15 years (83). The European Society of Radiology and the European Respiratory Society also recommend lung cancer screening in routine clinical practice at certified multidisciplinary medical centers (84). However, the increased number of CT scans is beyond the amount that radiologists can handle, and the high false-positive rate puts a strain on the lung cancer management system.

In this context, CAD could be an option for dealing with personnel shortages and false positives in lung cancer screening programs. There are two ways of utilizing CAD for lung cancer screening programs. The first is to mark the location of pulmonary nodules, and the second is to determine whether detected pulmonary nodules are malignant. In terms of nodule detection, CAD models before the use of DL techniques showed insufficient performance to be implemented in clinical practice (85-87). The performance gradually improved with the application of CNN-based DL models, and in 2016, the LUNg Nodule Analysis (LUNA) challenge was held for complete nodule detection and false-positive reduction based on 888 annotated images. The best model of the challenge achieved a sensitivity of 93% for the individual model and a sensitivity of 95% or more for the combined model (88).

Recent studies focus on determining whether a detected nodule is malignant or not beyond simple detection of pulmonary nodule. Ciompi *et al.* (37) developed a DL model that classified lung nodules into solid, non-solid, part-solid, calcified, periferic, and spiculated, which showed improved accuracy (39.9% *vs.* 79.5%) compared to conventional machine learning. Its accuracy was not inferior to that of six physician observers (69.6% *vs.* 72.9%). Ardila *et al.* (38) developed a model to predict the risk of lung cancer based on CT scans from the National Lung Cancer Screening Trial cases. The model had an AUC of 0.94, with an 11% reduction in false-positive rates and a 5% reduction in false-negative rates compared to radiologists when no prior CT images were provided. An example of a lung nodule identified on screening low-dose chest CT using an AI system is presented in *Figure 6*.

Diagnosis of COVID-19 in pandemic areas

Since its emergence in late 2019, COVID-19 has spread

worldwide, with more than 202 million confirmed patients and more than 4.2 million deaths worldwide as of August 8, 2021 (89). Healthcare workers have also been devoting more time and energy to COVID-19-related medical duties, facing shortages of equipment and supplies, as well as staffing shortages (90). To address these problems, researchers have rushed to develop AI models to support clinicians.

The primary diagnostic method for COVID-19 is the detection of SARS-CoV-2 via real-time reverse transcriptase-polymerase chain reaction (RT-PCR) in respiratory specimens. Although RT-PCR is the gold standard to diagnose COVID-19, imaging can complement its use to achieve greater diagnostic certainty and even serve as an alternative method in some regions where RT-PCR is not readily available. In some cases, CXR may exhibit findings of abnormalities in patients who initially had a negative RT-PCR test (91), and several recent studies have shown that chest CT has a higher sensitivity for COVID-19 than RT-PCR and can be considered as a screening tool for COVID-19 in pandemic areas (92-94).

Applying AI methods to COVID-19 radiologic imaging might enhance the accuracy of the diagnosis compared with RT-PCR, while also resolving the shortage of healthcare workers in pandemic areas. For example, AI can assist in the automated diagnosis and screening of COVID-19 using image analysis from CXR (95), CT scans (96-98) and lung ultrasonography (99). Harmon *et al.* (39) showed that a DL algorithm could achieve up to 90.8% accuracy, with 84% sensitivity and 93% specificity in the detection of COVID-19 pneumonia on chest CT using multinational datasets. In addition, some studies have suggested that AI can assist radiologists in distinguishing COVID-19 from other pulmonary infections on CXR (22) and chest CT (100,101). AI models have the potential to exploit the vast amount of multimodal data collected from patients and, if successful, could transform the detection, diagnosis, and triage of patients with suspected COVID-19. *Figure 7* shows an example of an AI-assisted interpretation of CXR with COVID-19-associated pneumonia.

Despite the potential of AI-assisted practice, there are still many limitations in deploying AI tools into the clinical workflow. Roberts *et al.* (102) found that none of the 415 papers selected for their study had a sufficiently documented manuscript describing a reproducible method, a methodology that followed best practices for developing AI models, and sufficient external validation. Further studies are required to address these issues before AI can take its

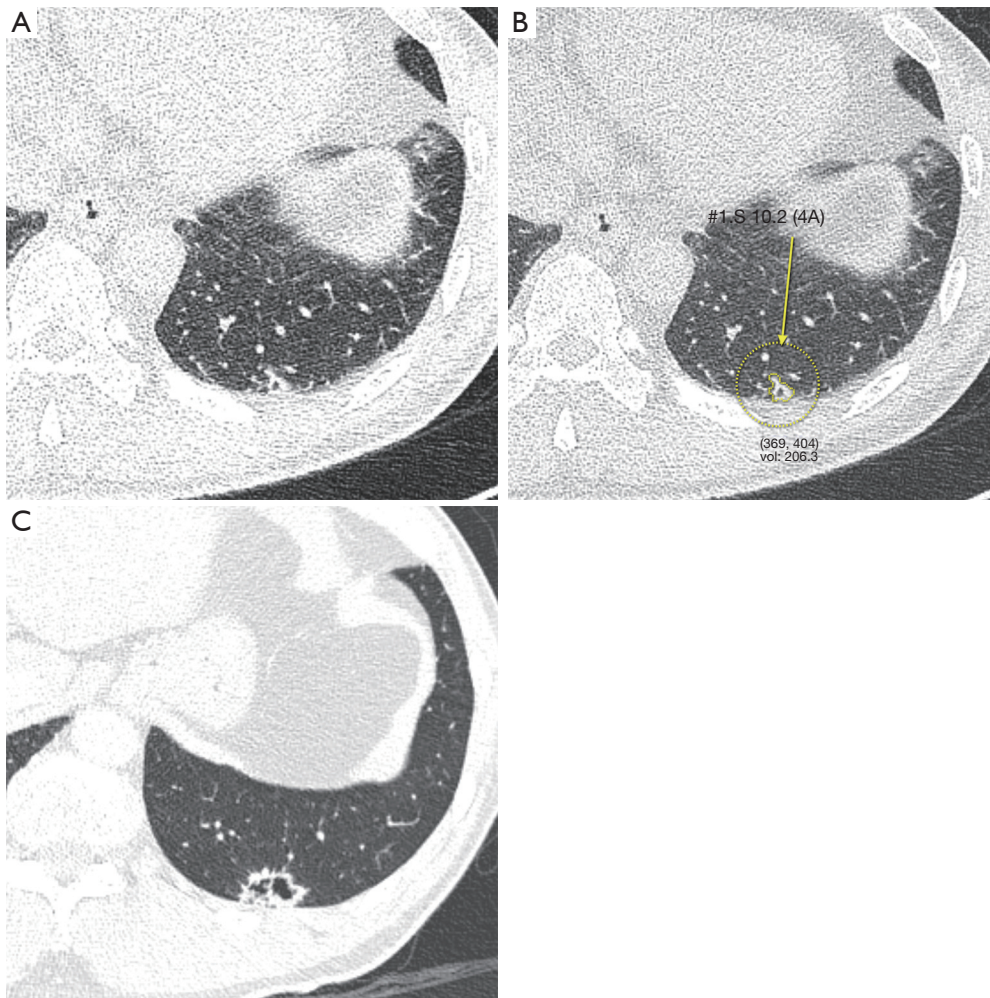


Figure 6 Identification and classification of a lung nodule on screening low-dose chest CT using an AI system. A screening low-dose chest CT scan of a 57-year-old ex-smoker (45 pack-years, quit smoking 7 years before) shows a small nodule with a cystic appearance at the left lower lobe of the lung (A). An AI system automatically identified the nodule. The average diameter of the nodule measured by the AI system was 10.2 mm, corresponding to Lung-RADS category 4A (B). A chest CT obtained 2 years later shows growth of the nodule, which was proven to be lung cancer (C).

place in the clinic.

Prognostic prediction and new discoveries

Most AI systems regarding thoracic diseases have focused on assisting the detection and diagnosis of radiologic abnormalities or diseases on imaging studies. Since the radiologic findings associated with specific thoracic diseases are well understood, AI systems are trained and evaluated to mimic the clinical practices of detecting disease-related findings. However, radiologists rarely know the outcome of patients undergoing radiologic examinations a decade

later; therefore, it is difficult to determine the imaging features with long-term prognostic value. DL algorithms can independently extract features from a large amount of data, and they have the potential to find novel imaging biomarkers. Thus, prognostication and therapeutic response prediction may be another important application of AI for thoracic diseases.

Early attempts to use AI for prognostic prediction were based on clinical information of patients (such as demographics, laboratory test results, treatment types, or gene expression data), and machine learning techniques have already shown superior performance over the existing

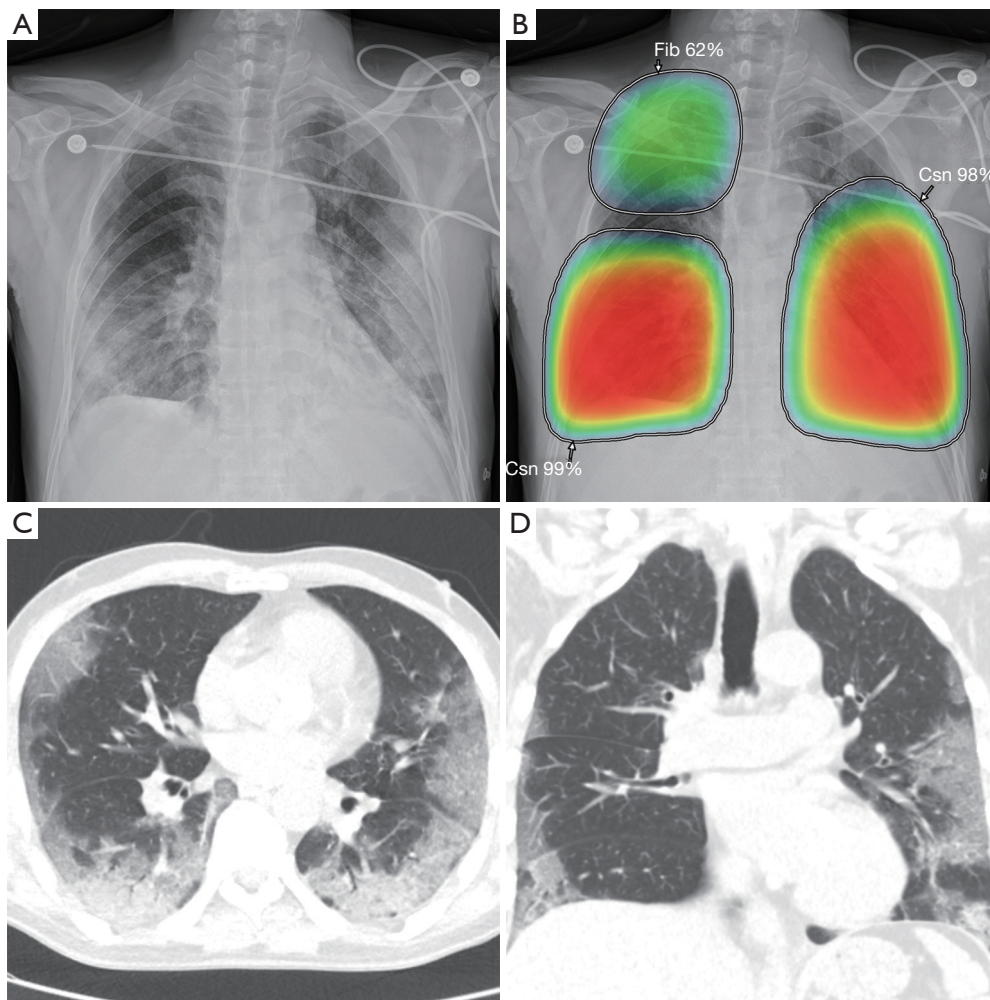


Figure 7 Identification of pneumonia associated with COVID-19 on chest radiograph using an AI CAD system. A chest radiograph of a 54-year-old male patient with COVID-19 shows diffusely increased opacities in both lung fields (A). The AI CAD system identified the opacities with a probability score of 99% (B). Chest CT of the patient shows regions of ground-glass opacities in both peripheral lungs, suggesting pneumonia associated with COVID-19 (C,D). COVID-19, coronavirus disease 2019; CAD, computer aided diagnosis.

survival prediction models (103,104). However, unlike mortality prediction models based on structured clinical information, the usage of image data for prognostic prediction is challenging. In this context, human-derived feature extractions from images can be advantageous, but the loss of important information during these procedures is unavoidable.

Hence, in 2017, González *et al.* (40) introduced a CNN-based DL model for acute respiratory disease event prediction (C-index, 0.64 and 0.55 for internal and external validation, respectively) and mortality prediction (C-index, 0.72 and 0.60, respectively) in smokers relying only on CT image data. To compare the performance of

conventional machine learning techniques with a DL-based prognostication model utilizing imaging data alone, Hosny *et al.* (41) presented a DL model for 2-year mortality prediction of non-small-cell lung cancer (NSCLC) patients. The model, which was trained and evaluated on seven independent NSCLC patient datasets across five institutions, outperformed existing structured data-based techniques, with AUCs of 0.70 and 0.71 for 2-year mortality after the start of radiotherapy and after surgery, respectively. Lu *et al.* (42) presented CXR-risk, a CNN-based DL model for 12-year mortality prediction from a single CXR. The high performance of CXR-risk proved the ability of CNN models to extract hidden prognostic

information from medical images. Additionally, in a recent study of Chao *et al.* (43), a CNN and Tri2D-Net based DL model successfully predicted cardiovascular mortality with LDCT images (AUC, 0.768), outperforming existing DL models.

DL can be used to find new discoveries for applications that are not part of current clinical practices. Raghu *et al.* (44) developed a DL algorithm that can estimate biological age (CXR-Age) from a chest radiograph to predict longevity beyond chronological age. Interestingly, their external validation tests performed on the PLCO and NLST populations showed significant improvements in the prediction of both long-term all-cause and cardiovascular mortality when CXR-age was used instead of chronological age (44).

Challenges and future directions of AI in pulmonary medicine

Applications of AI for thoracic diseases have demonstrated promising results in augmenting existing clinical systems, and prognostic prediction. However, there are still substantial limitations. For augmenting existing systems, as in the add-on scenario, it is not clear how to effectively integrate AI tools with physician decision-makers. Indeed, some studies showed no improvement of clinical outcomes with AI assistance in randomized controlled trials (105,106). The interaction between AI models and human users is poorly understood and little work has evaluated the potential impact of such systems. Interestingly, Gaube *et al.* (58) reported that radiologists rated advice as lower-quality when that advice seemed to come from an AI system, while non-radiology physicians with less task-expertise did not. Diagnostic accuracy was significantly worse when participants received inaccurate advice, regardless of the purported source. Their work raised the importance of the quality of advice and the importance of how advice (both from AI and non-AI sources) should be delivered in clinical environments.

Studies on DL models that can ingest multimodal data, including imaging data, clinical information and hopefully, genetic information, are also needed. Currently, the most advanced DL model for radiology applications only considers pixel value information, without data on the clinical background (107). However, in practice, relevant clinical information allows clinicians to interpret imaging results in the appropriate clinical context, providing information relevant for clinical decision-making and improving the diagnosis and prognosis. Apart from

the prognostic prediction of thoracic diseases, several DL models have demonstrated improved performance of prognostic prediction based on imaging data when complemented by other multimodal data. For example, in a study by Cheerla *et al.* (108), a pancancer survival prediction model integrating clinical, mRNA, miRNA, and whole-slide imaging data exhibited a C-index of 0.78. As asserted by Warth *et al.* (109), there exists a definite correlation between the morphological features of pathological images and genetic data in adenocarcinomas. This implies that integrating clinical and genetic data into image-based DL models may improve their performance.

Reproducibility and generalizability are also important issues. In many studies reporting excellent results, AI systems were only tested with internal validation data, which are retrospectively and non-rigorously collected (110). In this case, validation data may have an enriched disease prevalence and a narrow disease spectrum. In contrast, the population in real-world situations may have a much lower disease prevalence and a much broader spectrum of diseases, potentially hindering the performance of the DL algorithm. Thus, AI system needs to be further validated in real-world situations before they are implemented in clinical practice.

AI systems also need to be appropriately explained, with a particular focus on the logical background of their output. In order for a DL algorithm to receive credit or acceptance from physicians, it should appropriately explain the logical basis of the output (111). In particular, if a DL algorithm is used for prognostic prediction beyond existing clinical systems, clinicians may want to know why the algorithm provides certain outcomes based on their existing knowledge. AI systems should provide interpretability to receive credibility from physicians and be implemented in clinical practice.

Conclusions

AI has shown considerable potential for many thoracic diseases, particularly in the field of thoracic radiology. This promising technique is expected to effectively address various clinical problems that have not been solved due to a lack of clinical resources or technological limitations. AI could be a cost-effective and excellent second reader, providing automated quantification and prioritization. In addition, AI models could be used to improve screening for tuberculosis and lung cancer, as well as for prognostic prediction. It is necessary not only to advance the performance of AI systems, but also to understand and

discuss how to use AI in clinical practice. With advances in technology and appropriate preparation of physicians, AI could transform current medical practices and contribute to improving human health.

Acknowledgments

Funding: This study was supported by the Seoul National University Hospital Research Fund (Grant No. 03-2021-0270).

Footnote

Provenance and Peer Review: This article was commissioned by the Guest Editors (Jianxing He and Hengrui Liang) for the series “Artificial Intelligence in Thoracic Disease: from Bench to Bed” published in *Journal of Thoracic Disease*. The article has undergone external peer review.

Reporting Checklist: The authors have completed the Narrative Review reporting checklist. Available at <https://dx.doi.org/10.21037/jtd-21-1342>

Peer Review File: Available at <https://dx.doi.org/10.21037/jtd-21-1342>

Conflicts of Interest: The authors have completed the ICMJE uniform disclosure form (available at <https://dx.doi.org/10.21037/jtd-21-1342>). The series “Artificial Intelligence in Thoracic Disease: from Bench to Bed” was commissioned by the editorial office without any funding or sponsorship. EJH reports research grants from Lunit, Coreline Soft, and Monitor Corporation, outside the present study. CMP reports that this study was supported by research fund from Seoul National University Hospital. CMP is a committee chair of Radiology Investigators Network of Korea, committee chair of International Liaison, and board member of Korean Society of Artificial Intelligence in Medicine. CMP holds stock of Promedius and stock options of Coreline soft and Lunit. The authors have no other conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Open Access Statement: This is an Open Access article

distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

1. Gulshan V, Peng L, Coram M, et al. Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. *JAMA* 2016;316:2402-10.
2. Ting DSW, Cheung CY, Lim G, et al. Development and Validation of a Deep Learning System for Diabetic Retinopathy and Related Eye Diseases Using Retinal Images From Multiethnic Populations With Diabetes. *JAMA* 2017;318:2211-23.
3. Poplin R, Varadarajan AV, Blumer K, et al. Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning. *Nat Biomed Eng* 2018;2:158-64.
4. De Fauw J, Ledsam JR, Romera-Paredes B, et al. Clinically applicable deep learning for diagnosis and referral in retinal disease. *Nat Med* 2018;24:1342-50.
5. Ghorbani A, Ouyang D, Abid A, et al. Deep learning interpretation of echocardiograms. *NPJ Digit Med* 2020;3:10.
6. Ulloa Cerna AE, Jing L, Good CW, et al. Deep-learning-assisted analysis of echocardiographic videos improves predictions of all-cause mortality. *Nat Biomed Eng* 2021;5:546-54.
7. Ouyang D, He B, Ghorbani A, et al. Video-based AI for beat-to-beat assessment of cardiac function. *Nature* 2020;580:252-6.
8. McKinney SM, Sieniek M, Godbole V, et al. International evaluation of an AI system for breast cancer screening. *Nature* 2020;577:89-94.
9. Hwang EJ, Kim H, Lee JH, et al. Automated identification of chest radiographs with referable abnormality with deep learning: need for recalibration. *Eur Radiol* 2020;30:6902-12.
10. Nam JG, Kim J, Noh K, et al. Automatic prediction of left cardiac chamber enlargement from chest radiographs using convolutional neural network. *Eur Radiol* 2021;31:8130-40.
11. Hwang EJ, Park CM. Clinical Implementation of Deep Learning in Thoracic Radiology: Potential Applications and Challenges. *Korean J Radiol* 2020;21:511-25.

12. Kim JH, Kim JY, Kim GH, et al. Clinical Validation of a Deep Learning Algorithm for Detection of Pneumonia on Chest Radiographs in Emergency Department Patients with Acute Febrile Respiratory Illness. *J Clin Med* 2020;9:1981.
13. Hwang EJ, Kim KB, Kim JY, et al. COVID-19 pneumonia on chest X-rays: Performance of a deep learning-based computer-aided detection system. *PLoS One* 2021;16:e0252440.
14. Kim H, Lee D, Cho WS, et al. CT-based deep learning model to differentiate invasive pulmonary adenocarcinomas appearing as subsolid nodules among surgical candidates: comparison of the diagnostic performance with a size-based logistic model and radiologists. *Eur Radiol* 2020;30:3295-305.
15. Hwang EJ, Hong JH, Lee KH, et al. Deep learning algorithm for surveillance of pneumothorax after lung biopsy: a multicenter diagnostic cohort study. *Eur Radiol* 2020;30:3660-71.
16. Hwang EJ, Nam JG, Lim WH, et al. Deep Learning for Chest Radiograph Diagnosis in the Emergency Department. *Radiology* 2019;293:573-80.
17. Lee JH, Park S, Hwang EJ, et al. Deep learning-based automated detection algorithm for active pulmonary tuberculosis on chest radiographs: diagnostic performance in systematic screening of asymptomatic individuals. *Eur Radiol* 2021;31:1069-80.
18. Nam JG, Kim M, Park J, et al. Development and validation of a deep learning algorithm detecting 10 common abnormalities on chest radiographs. *Eur Respir J* 2021;57:2003061.
19. Hwang EJ, Park S, Jin KN, et al. Development and Validation of a Deep Learning-Based Automated Detection Algorithm for Major Thoracic Diseases on Chest Radiographs. *JAMA Netw Open* 2019;2:e191095.
20. Hwang EJ, Park S, Jin KN, et al. Development and Validation of a Deep Learning-based Automatic Detection Algorithm for Active Pulmonary Tuberculosis on Chest Radiographs. *Clin Infect Dis* 2019;69:739-47.
21. Nam JG, Park S, Hwang EJ, et al. Development and Validation of Deep Learning-based Automatic Detection Algorithm for Malignant Pulmonary Nodules on Chest Radiographs. *Radiology* 2019;290:218-28.
22. Hwang EJ, Kim H, Yoon SH, et al. 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.
23. Lee JH, Sun HY, Park S, et al. Performance of a Deep Learning Algorithm Compared with Radiologic Interpretation for Lung Cancer Detection on Chest Radiographs in a Health Screening Population. *Radiology* 2020;297:687-96.
24. Choi H, Kim H, Hong W, et al. Prediction of visceral pleural invasion in lung cancer on CT: deep learning model achieves a radiologist-level performance with adaptive sensitivity and specificity to clinical needs. *Eur Radiol* 2021;31:2866-76.
25. Kim H, Goo JM, Lee KH, et al. Preoperative CT-based Deep Learning Model for Predicting Disease-Free Survival in Patients with Lung Adenocarcinomas. *Radiology* 2020;296:216-24.
26. Kim H, Park CM, Goo JM. Test-retest reproducibility of a deep learning-based automatic detection algorithm for the chest radiograph. *Eur Radiol* 2020;30:2346-55.
27. Nam JG, Hwang EJ, Kim DS, et al. Undetected Lung Cancer at Posteroanterior Chest Radiography: Potential Role of a Deep Learning-based Detection Algorithm. *Radiol Cardiothorac Imaging* 2020;2:e190222.
28. Ferlay J, Colombet M, Soerjomataram I, et al. Cancer statistics for the year 2020: An overview. *Int J Cancer* 2021. [Epub ahead of print]. doi: 10.1002/ijc.33588.
29. Chakaya J, Khan M, Ntoumi F, et al. Global Tuberculosis Report 2020 - Reflections on the Global TB burden, treatment and prevention efforts. *Int J Infect Dis* 2021;113:S7-12.
30. UNSCEAR. Sources and effects of ionizing radiation. United Nations New York; 2000.
31. Seah JCY, Tang CHM, Buchlak QD, et al. Effect of a comprehensive deep-learning model on the accuracy of chest x-ray interpretation by radiologists: a retrospective, multireader multicase study. *Lancet Digit Health* 2021;3:e496-506.
32. Wu E, Wu K, Daneshjou R, et al. How medical AI devices are evaluated: limitations and recommendations from an analysis of FDA approvals. *Nat Med* 2021;27:582-4.
33. Huang SC, Kothari T, Banerjee I, et al. PENet-a scalable deep-learning model for automated diagnosis of pulmonary embolism using volumetric CT imaging. *NPJ Digit Med* 2020;3:61.
34. Hata A, Yanagawa M, Yamagata K, et al. Deep learning algorithm for detection of aortic dissection on non-contrast-enhanced CT. *Eur Radiol* 2021;31:1151-9.
35. Hasenstab KA, Yuan N, Retson T, et al. Automated CT Staging of Chronic Obstructive Pulmonary Disease Severity for Predicting Disease Progression and Mortality with a Deep Learning Convolutional Neural Network.

- Radiol Cardiothorac Imaging 2021;3:e200477.
36. Chassagnon G, Vakilopoulou M, Régent A, et al. Deep Learning-based Approach for Automated Assessment of Interstitial Lung Disease in Systemic Sclerosis on CT Images. *Radiol Artif Intell* 2020;2:e190006.
 37. Ciompi F, Chung K, van Riel SJ, et al. Towards automatic pulmonary nodule management in lung cancer screening with deep learning. *Sci Rep* 2017;7:46479.
 38. Ardila D, Kiraly AP, Bharadwaj S, et al. End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. *Nat Med* 2019;25:954-61.
 39. Harmon SA, Sanford TH, Xu S, et al. Artificial intelligence for the detection of COVID-19 pneumonia on chest CT using multinational datasets. *Nat Commun* 2020;11:4080.
 40. González G, Ash SY, Vegas-Sánchez-Ferrero G, et al. Disease Staging and Prognosis in Smokers Using Deep Learning in Chest Computed Tomography. *Am J Respir Crit Care Med* 2018;197:193-203.
 41. Hosny A, Parmar C, Coroller TP, et al. Deep learning for lung cancer prognostication: A retrospective multi-cohort radiomics study. *PLoS Med* 2018;15:e1002711.
 42. Lu MT, Ivanov A, Mayrhofer T, et al. Deep Learning to Assess Long-term Mortality From Chest Radiographs. *JAMA Netw Open* 2019;2:e197416.
 43. Chao H, Shan H, Homayounieh F, et al. Deep learning predicts cardiovascular disease risks from lung cancer screening low dose computed tomography. *Nat Commun* 2021;12:2963.
 44. Raghu VK, Weiss J, Hoffmann U, et al. Deep Learning to Estimate Biological Age From Chest Radiographs. *JACC Cardiovasc Imaging* 2021;14:2226-36.
 45. National Center for Health Statistics. National Hospital Ambulatory Medical Care Survey: 2017 emergency department summary tables. Available online: https://www.cdc.gov/nchs/data/nhamcs/web_tables/2017_ed_web_tables-508.pdf
 46. National Center for Health Statistics. National Hospital Ambulatory Medical Care Survey: 2011 Emergency Department Summary Tables. Available online: https://www.cdc.gov/nchs/data/ahcd/nhamcs_emergency/2011_ed_web_tables.pdf
 47. National Center for Health Statistics. National Hospital Ambulatory Medical Care Survey: 2015 Emergency Department Summary Tables. Available online: https://www.cdc.gov/nchs/data/nhamcs/web_tables/2015_ed_web_tables.pdf
 48. Chung JH, Duszak R Jr, Hemingway J, et al. Increasing Utilization of Chest Imaging in US Emergency Departments From 1994 to 2015. *J Am Coll Radiol* 2019;16:674-82.
 49. Prologo JD, Gilkeson RC, Diaz M, et al. CT pulmonary angiography: a comparative analysis of the utilization patterns in emergency department and hospitalized patients between 1998 and 2003. *AJR Am J Roentgenol* 2004;183:1093-6.
 50. Chandra S, Sarkar PK, Chandra D, et al. Finding an alternative diagnosis does not justify increased use of CT-pulmonary angiography. *BMC Pulm Med* 2013;13:9.
 51. Braverman AC. Acute aortic dissection: clinician update. *Circulation* 2010;122:184-8.
 52. Brunswick JE, Ilkhanipour K, Seaberg DC, et al. Radiographic interpretation in the emergency department. *Am J Emerg Med* 1996;14:346-8.
 53. Mayhue FE, Rust DD, Aldag JC, et al. Accuracy of interpretations of emergency department radiographs: effect of confidence levels. *Ann Emerg Med* 1989;18:826-30.
 54. Preston CA, Marr JJ 3rd, Amaraneni KK, et al. Reduction of "callbacks" to the ED due to discrepancies in plain radiograph interpretation. *Am J Emerg Med* 1998;16:160-2.
 55. Gatt ME, Spectre G, Paltiel O, et al. Chest radiographs in the emergency department: is the radiologist really necessary? *Postgrad Med J* 2003;79:214-7.
 56. Geijer H, Geijer M. Added value of double reading in diagnostic radiology, a systematic review. *Insights Imaging* 2018;9:287-301.
 57. Gonem S, Janssens W, Das N, et al. Applications of artificial intelligence and machine learning in respiratory medicine. *Thorax* 2020;75:695-701.
 58. Gaube S, Suresh H, Raue M, et al. Do as AI say: susceptibility in deployment of clinical decision-aids. *NPJ Digit Med* 2021;4:31.
 59. Lynch DA, Al-Qaisi MA. Quantitative computed tomography in chronic obstructive pulmonary disease. *J Thorac Imaging* 2013;28:284-90.
 60. Nambu A, Zach J, Schroeder J, et al. Quantitative computed tomography measurements to evaluate airway disease in chronic obstructive pulmonary disease: Relationship to physiological measurements, clinical index and visual assessment of airway disease. *Eur J Radiol* 2016;85:2144-51.
 61. Coxson HO, Rogers RM. Quantitative computed tomography of chronic obstructive pulmonary disease. *Acad Radiol* 2005;12:1457-63.
 62. Bartholmai BJ, Raghunath S, Karwoski RA, et al. Quantitative computed tomography imaging of interstitial

- lung diseases. *J Thorac Imaging* 2013;28:298-307.
63. Dong X, Lei Y, Wang T, et al. Automatic multiorgan segmentation in thorax CT images using U-net-GAN. *Med Phys* 2019;46:2157-68.
 64. Park B, Park H, Lee SM, et al. Lung Segmentation on HRCT and Volumetric CT for Diffuse Interstitial Lung Disease Using Deep Convolutional Neural Networks. *J Digit Imaging* 2019;32:1019-26.
 65. Park J, Yun J, Kim N, et al. Fully Automated Lung Lobe Segmentation in Volumetric Chest CT with 3D U-Net: Validation with Intra- and Extra-Datasets. *J Digit Imaging* 2020;33:221-30.
 66. Yun J, Park J, Yu D, et al. Improvement of fully automated airway segmentation on volumetric computed tomographic images using a 2.5 dimensional convolutional neural net. *Med Image Anal* 2019;51:13-20.
 67. Aresta G, Jacobs C, Araújo T, et al. iW-Net: an automatic and minimalistic interactive lung nodule segmentation deep network. *Sci Rep* 2019;9:11591.
 68. Wang S, Zhou M, Liu Z, et al. Central focused convolutional neural networks: Developing a data-driven model for lung nodule segmentation. *Med Image Anal* 2017;40:172-83.
 69. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More than Pictures, They Are Data. *Radiology* 2016;278:563-77.
 70. Gevaert O, Mitchell LA, Achrol AS, et al. Glioblastoma multiforme: exploratory radiogenomic analysis by using quantitative image features. *Radiology* 2014;273:168-74.
 71. Aerts HJ, Velazquez ER, Leijenaar RT, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat Commun* 2014;5:4006.
 72. Berenguer R, Pastor-Juan MDR, Canales-Vázquez J, et al. Radiomics of CT Features May Be Nonreproducible and Redundant: Influence of CT Acquisition Parameters. *Radiology* 2018;288:407-15.
 73. Choe J, Lee SM, Do KH, et al. Deep Learning-based Image Conversion of CT Reconstruction Kernels Improves Radiomics Reproducibility for Pulmonary Nodules or Masses. *Radiology* 2019;292:365-73.
 74. World Health Organization. Systematic screening for active tuberculosis: principles and recommendations. World Health Organization; 2013.
 75. World Health Organization. WHO operational handbook on tuberculosis. Module 2: Screening - Systematic screening for tuberculosis disease. Geneva: 2021.
 76. Kumar N, Bhargava SK, Agrawal CS, et al. Chest radiographs and their reliability in the diagnosis of tuberculosis. *JNMA J Nepal Med Assoc* 2005;44:138-42.
 77. Saktiawati AMI, Subronto YW, Stienstra Y, et al. Sensitivity and specificity of routine diagnostic work-up for tuberculosis in lung clinics in Yogyakarta, Indonesia: a cohort study. *BMC Public Health* 2019;19:363.
 78. van't Hoog AH, Meme HK, Laserson KF, et al. Screening strategies for tuberculosis prevalence surveys: the value of chest radiography and symptoms. *PLoS One* 2012;7:e38691.
 79. Hwang S, Kim HE, Jeong J, et al. A novel approach for tuberculosis screening based on deep convolutional neural networks. *Proc. SPIE 9785, Medical Imaging 2016: Computer-Aided Diagnosis, 97852W* (24 March 2016). doi: 10.1117/12.2216198.
 80. Qin ZZ, Sander MS, Rai B, et al. Using artificial intelligence to read chest radiographs for tuberculosis detection: A multi-site evaluation of the diagnostic accuracy of three deep learning systems. *Sci Rep* 2019;9:15000.
 81. Khan FA, Majidulla A, Tavaziva G, et al. Chest x-ray analysis with deep learning-based software as a triage test for pulmonary tuberculosis: a prospective study of diagnostic accuracy for culture-confirmed disease. *Lancet Digit Health* 2020;2:e573-81.
 82. National Lung Screening Trial Research Team; Aberle DR, Adams AM, et al. Reduced lung-cancer mortality with low-dose computed tomographic screening. *N Engl J Med* 2011;365:395-409.
 83. US Preventive Services Task Force; Krist AH, Davidson KW, et al. Screening for Lung Cancer: US Preventive Services Task Force Recommendation Statement. *JAMA* 2021;325:962-70.
 84. Kauczor HU, Bonomo L, Gaga M, et al. ESR/ERS white paper on lung cancer screening. *Eur Radiol* 2015;25:2519-31.
 85. Armato SG 3rd, Altman MB, La Rivière PJ. Automated detection of lung nodules in CT scans: effect of image reconstruction algorithm. *Med Phys* 2003;30:461-72.
 86. Suzuki K, Armato SG 3rd, Li F, et al. Massive training artificial neural network (MTANN) for reduction of false positives in computerized detection of lung nodules in low dose computed tomography. *Med Phys* 2003;30:1602-17.
 87. Roos JE, Paik D, Olsen D, et al. Computer-aided detection (CAD) of lung nodules in CT scans: radiologist performance and reading time with incremental CAD assistance. *Eur Radiol* 2010;20:549-57.
 88. Setio AAA, Traverso A, de Bel T, et al. Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography

- images: The LUNA16 challenge. *Med Image Anal* 2017;42:1-13.
89. World Health Organization. Weekly epidemiological update on COVID-19 - 10 August 2021. Available online: <https://www.who.int/publications/m/item/weekly-epidemiological-update-on-covid-19---10-august-2021>
 90. El Haj M, Allain P, Annweiler C, et al. Burnout of Healthcare Workers in Acute Care Geriatric Facilities During the COVID-19 Crisis: An Online-Based Study. *J Alzheimers Dis* 2020;78:847-52.
 91. Wong HYF, Lam HYS, Fong AH, et al. Frequency and Distribution of Chest Radiographic Findings in Patients Positive for COVID-19. *Radiology* 2020;296:E72-8.
 92. Fang Y, Zhang H, Xie J, et al. Sensitivity of Chest CT for COVID-19: Comparison to RT-PCR. *Radiology* 2020;296:E115-7.
 93. Long C, Xu H, Shen Q, et al. Diagnosis of the Coronavirus disease (COVID-19): rRT-PCR or CT? *Eur J Radiol* 2020;126:108961.
 94. Ai T, Yang Z, Hou H, et al. Correlation of Chest CT and RT-PCR Testing for Coronavirus Disease 2019 (COVID-19) in China: A Report of 1014 Cases. *Radiology* 2020;296:E32-40.
 95. Ozturk T, Talo M, Yildirim EA, et al. Automated detection of COVID-19 cases using deep neural networks with X-ray images. *Comput Biol Med* 2020;121:103792.
 96. Majidi H, Niksolat F. Chest CT in patients suspected of COVID-19 infection: A reliable alternative for RT-PCR. *Am J Emerg Med* 2020;38:2730-2.
 97. Wang S, Kang B, Ma J, et al. A deep learning algorithm using CT images to screen for Corona virus disease (COVID-19). *Eur Radiol* 2021;31:6096-104.
 98. Li Z, Zhong Z, Li Y, et al. From community-acquired pneumonia to COVID-19: a deep learning-based method for quantitative analysis of COVID-19 on thick-section CT scans. *Eur Radiol* 2020;30:6828-37.
 99. Roy S, Menapace W, Oei S, et al. Deep Learning for Classification and Localization of COVID-19 Markers in Point-of-Care Lung Ultrasound. *IEEE Trans Med Imaging* 2020;39:2676-87.
 100. Bai HX, Wang R, Xiong Z, et al. Artificial Intelligence Augmentation of Radiologist Performance in Distinguishing COVID-19 from Pneumonia of Other Origin at Chest CT. *Radiology* 2020;296:E156-65.
 101. Yang Y, Lure FYM, Miao H, et al. Using artificial intelligence to assist radiologists in distinguishing COVID-19 from other pulmonary infections. *J Xray Sci Technol* 2021;29:1-17.
 102. Roberts M, Driggs D, Thorpe M, et al. Common pitfalls and recommendations for using machine learning to detect and prognosticate for COVID-19 using chest radiographs and CT scans. *Nat Mach Intell* 2021;3:199-217.
 103. Matsuo K, Purushotham S, Jiang B, et al. Survival outcome prediction in cervical cancer: Cox models vs deep-learning model. *Am J Obstet Gynecol* 2019;220:381.e1-381.e14.
 104. Katzman JL, Shaham U, Cloninger A, et al. DeepSurv: personalized treatment recommender system using a Cox proportional hazards deep neural network. *BMC Med Res Methodol* 2018;18:24.
 105. INFANT Collaborative Group. Computerised interpretation of fetal heart rate during labour (INFANT): a randomised controlled trial. *Lancet* 2017;389:1719-29.
 106. Nunes I, Ayres-de-Campos D, Ugwumadu A, et al. Central Fetal Monitoring With and Without Computer Analysis: A Randomized Controlled Trial. *Obstet Gynecol* 2017;129:83-90.
 107. Huang SC, Pareek A, Seyyedi S, et al. Fusion of medical imaging and electronic health records using deep learning: a systematic review and implementation guidelines. *NPJ Digit Med* 2020;3:136.
 108. Cheerla A, Gevaert O. Deep learning with multimodal representation for pancancer prognosis prediction. *Bioinformatics* 2019;35:i446-54.
 109. Warth A, Muley T, Meister M, et al. The novel histologic International Association for the Study of Lung Cancer/American Thoracic Society/European Respiratory Society classification system of lung adenocarcinoma is a stage-independent predictor of survival. *J Clin Oncol* 2012;30:1438-46.
 110. Kim DW, Jang HY, Kim KW, et al. Design Characteristics of Studies Reporting the Performance of Artificial Intelligence Algorithms for Diagnostic Analysis of Medical Images: Results from Recently Published Papers. *Korean J Radiol* 2019;20:405-10.
 111. Choo J, Liu S. Visual Analytics for Explainable Deep Learning. *IEEE Comput Graph Appl* 2018;38:84-92.

Cite this article as: Kim Y, Park JY, Hwang EJ, Lee SM, Park CM. Applications of artificial intelligence in the thorax: a narrative review focusing on thoracic radiology. *J Thorac Dis* 2021;13(12):6943-6962. doi: 10.21037/jtd-21-1342