



New trend in artificial intelligence-based assistive technology for thoracic imaging

Masahiro Yanagawa¹ · Rintaro Ito² · Taiki Nozaki³ · Tomoyuki Fujioka⁴ · Akira Yamada⁵ · Shohei Fujita⁶ · Koji Kamagata⁷ · Yasutaka Fushimi⁸ · Takahiro Tsuboyama¹ · Yusuke Matsui⁹ · Fuminari Tatsugami¹⁰ · Mariko Kawamura² · Daiju Ueda¹¹ · Noriyuki Fujima¹² · Takeshi Nakaura¹³ · Kenji Hirata¹⁴ · Shinji Naganawa²

Received: 2 July 2023 / Accepted: 25 July 2023 / Published online: 28 August 2023
© The Author(s) 2023

Abstract

Although there is no solid agreement for artificial intelligence (AI), it refers to a computer system with intelligence similar to that of humans. Deep learning appeared in 2006, and more than 10 years have passed since the third AI boom was triggered by improvements in computing power, algorithm development, and the use of big data. In recent years, the application and development of AI technology in the medical field have intensified internationally. There is no doubt that AI will be used in clinical practice to assist in diagnostic imaging in the future. In qualitative diagnosis, it is desirable to develop an explainable AI that at least represents the basis of the diagnostic process. However, it must be kept in mind that AI is a physician-assistant system, and the final decision should be made by the physician while understanding the limitations of AI. The aim of this article is to review the application of AI technology in diagnostic imaging from PubMed database while particularly focusing on diagnostic imaging in thorax such as lesion detection and qualitative diagnosis in order to help radiologists and clinicians to become more familiar with AI in thorax.

Keywords Artificial intelligence · Deep learning · Convolutional neural network · Vision transformer · Explainable AI · Thoracic imaging

✉ Masahiro Yanagawa
m-yanagawa@radiol.med.osaka-u.ac.jp

¹ Department of Radiology, Osaka University Graduate School of Medicine, 2-2 Yamadaoka, Suita-City, Osaka 565-0871, Japan

² Department of Radiology, Nagoya University Graduate School of Medicine, 65 Tsurumai-cho, Showa-ku, Nagoya, Aichi 466-8550, Japan

³ Department of Radiology, Keio University School of Medicine, 35 Shinanomachi, Shinjuku-ku, Tokyo 160-0016, Japan

⁴ Department of Diagnostic Radiology, Tokyo Medical and Dental University, 1-5-45 Yushima, Bunkyo-ku, Tokyo 113-8519, Japan

⁵ Department of Radiology, Shinshu University School of Medicine, 3-1-1 Asahi, Matsumoto, Nagano 390-2621, Japan

⁶ Department of Radiology, Graduate School of Medicine and Faculty of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8655, Japan

⁷ Department of Radiology, Juntendo University Graduate School of Medicine, Bunkyo-ku, Tokyo 113-8421, Japan

⁸ Department of Diagnostic Imaging and Nuclear Medicine, Kyoto University Graduate School of Medicine, 54 Shogoin Kawaharacho, Sakyo-ku, Kyoto 606-8507, Japan

⁹ Department of Radiology, Faculty of Medicine, Dentistry and Pharmaceutical Sciences, Okayama University, 2-5-1 Shikata-cho, Kita-ku, Okayama 700-8558, Japan

¹⁰ Department of Diagnostic Radiology, Hiroshima University, 1-2-3 Kasumi, Minami-ku, Hiroshima 734-8551, Japan

¹¹ Department of Diagnostic and Interventional Radiology, Graduate School of Medicine, Osaka Metropolitan University, 1-4-3 Asahi-Machi, Abeno-ku, Osaka 545-8585, Japan

¹² Department of Diagnostic and Interventional Radiology, Hokkaido University Hospital, N15, W5, Kita-ku, Sapporo 060-8638, Japan

¹³ Department of Diagnostic Radiology, Kumamoto University Graduate School of Medicine, 1-1-1 Honjo Chuo-ku, Kumamoto 860-8556, Japan

¹⁴ Department of Diagnostic Imaging, Graduate School of Medicine, Hokkaido University, Kita 15 Nish 17, Kita-ku, Sapporo, Hokkaido 060-8648, Japan

Introduction

Looking back on the history of artificial intelligence (AI) [1–3], the first AI boom emerged in the 1950s, when computational reasoning and exploration flourished and expectations for AI grew. The next boom was the second AI boom in the 1980s. A system that responds with conditioned reflexes by teaching AI knowledge in the form of rules, called an expert system, was put into practical use. Machine learning is a technique in which a program learns from data, extracts patterns, and makes predictions and judgments on unknown data. Typical algorithms include support vector machines, random forests, and decision trees. Deep learning is a branch of machine learning based on hierarchical and multilayer artificial neural networks. Deep learning appeared in 2006, and it has been more than 10 years since the third AI boom was triggered by improvements in computing power, algorithm development, and the use of big data. Recent advancements in AI, particularly in deep learning, have revolutionized the field of medical image analysis. In the past, medical image analysis relied heavily on systems created by human domain experts, utilizing statistical or machine learning models that required manual selection of image features or regions of interest [4]. However, recent deep learning models can automatically learn these image features with minimal human intervention, resulting in more efficient and resource-saving image analysis tasks [5]. Currently, AI techniques have been developed for various fields of medical imaging including chest radiography, computed tomography (CT), magnetic resonance imaging (MRI), positron emission tomography (PET-CT), and ultrasonography [6–23]. In the thorax, particularly, AI has already been applied in clinical settings, such as nodule detection support, benign/malignant diagnosis support, image processing such as automatic quantification, and image quality improvement through noise reduction processing.

This article mainly outlines detection, diagnosis, and prognosis, focusing on image recognition, lung cancer as a representative lung tumor, coronavirus disease 2019 (COVID-19) as an infectious disease, and pulmonary embolism (PE) as a non-tumor/non-infection. This review can help readers understand the recent trends in AI in the field of thoracic imaging area and how to use AI as it evolves.

Materials and methods

For the literature survey, PubMed database up to April 30, 2023, was used to search studies relevant to evaluation of clinical applications in thorax using AI. Keywords

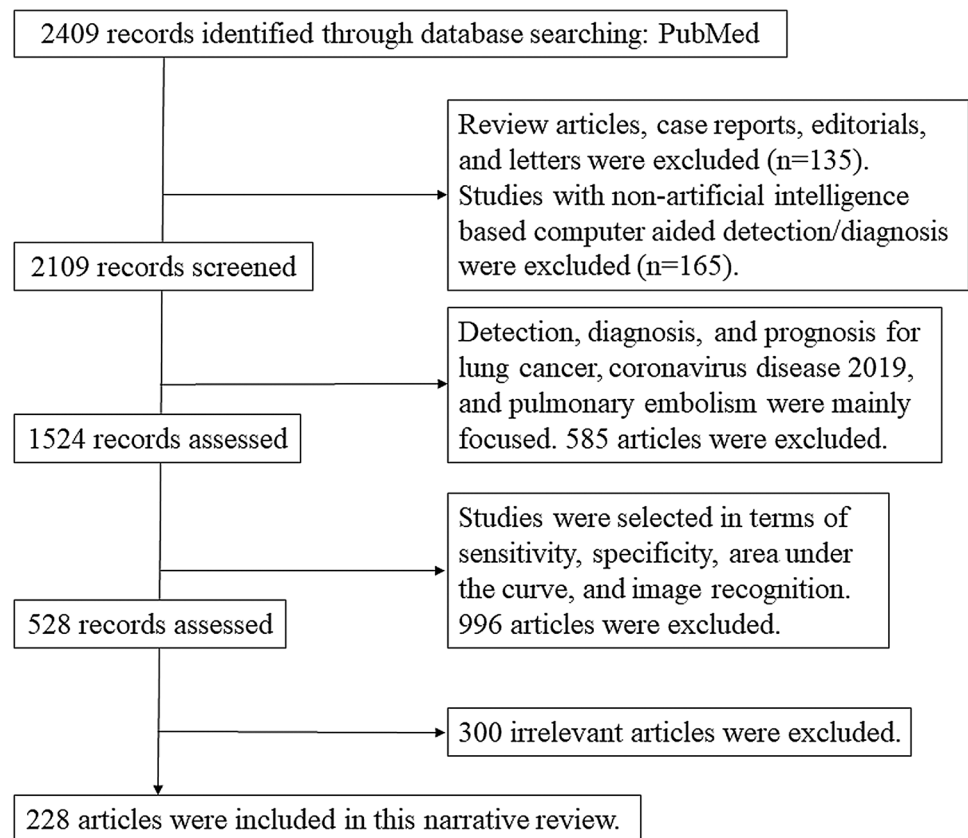
searched included the following: {lung imaging “English” [Language] AND (“artificial intelligence” OR “deep learning” OR “neural network” OR “machine learning” OR “computer aided”) AND (“CT”) NOT (“ultrasonography” OR “PET” OR “SPECT”)}. In particular, detection, diagnosis, and prognosis for lung cancer, COVID-19, and PE were mainly focused in this review. Citations and references from the retrieved studies were used as additional sources for literature review, and manually search was conducted. Review articles, case reports, editorials, and letters were excluded. Exclusion criteria were as follows: (a) researches outside the subject area; (b) phantom or animal or healthy individual researches; (c) researches focused on methodological aspects alone; (d) researches by a variety of methods; (e) researches on organs other than the lung; and (f) researches without evaluation of clinical outcomes. Figure 1 shows the flowchart of the selection process. Finally, 228 articles were included in the narrative review.

AI mechanisms commonly used for image recognition

The model commonly used in image recognition is a convolutional neural network (CNN), in which the convolution and pooling layers are repeated to produce the output, while performing image pattern matching, data compression, and misalignment correction. The CNNs have been used in many AI researches [6–22], which have shown promising performance in detecting and diagnosing diseases on medical imaging, sometimes outperforming radiologists [7, 19, 24]. However, one of the problems of AI is overfitting [25]. That is, AI works for known training data but does not fit for unknown data due to limitations caused by the generalizability of the variables under the shift of the data distribution. In addition, the high degree of freedom in the design of hidden layers is a factor in the excellent expressiveness of neural networks, but hidden stratification may cause better or worse performance of CNN models on certain subsets [26, 27]. CNN models provide excellent output but leave room for improvement.

As a relatively new image recognition model, a model called vision transformer (ViT), which does not use the convolution operation, has appeared [28, 29] and has shown promising performance on general image datasets. The ViT model processes images not as pixels but as patches of a certain size by applying techniques used in natural language processing to image processing, which uses an analogous approach by dividing an input image into some pixel patches and projecting them into a linear embedding. The sequence is input to the transformer encoder, which contains multi-headed self-attention layers and multilayer perceptron blocks after layer normalization to reduce training time and improve

Fig. 1 Flowchart of studies selection process



generalization performance. The transformer encoder outputs feature vectors corresponding to the input patches, followed by softmax activation. Due to the parallel processing by multi-headed self-attention layers and multilayer perceptron blocks, there is an advantage in terms of computational cost, and there is also an advantage that performance does not degrade even when the model is large. Recently, Murphy et al. [30] compared that ViT and CNN for disease classification on chest radiographs. CNN models consistently outperformed ViT models, although the absolute differences in performance were small and may not be clinically significant. On the other hand, ViT was less susceptible to specific kinds of hidden stratification known to affect CNN models. At this time, it may be too early for ViT to be a replacement for CNN in the diagnosis of disease on chest radiographs. Although ViTs have shown improvements over CNNs in general imaging datasets [29, 31], the usefulness of ViT must be further examined in the future.

Clinical application to lung tumors: detectability of pulmonary nodules

The leading cause of cancer-related deaths is lung cancer [32], and in our country, the mainstream approach to reducing lung cancer mortality is screening with chest radiography. The results of two randomized controlled trials, the

National Lung Screening Trial (NLST) in the USA in 2011 [33] and the Dutch-Belgian randomized lung cancer screening trial (NELSON) in Europe in 2020 [34], suggest that low-dose CT screening may be effective in reducing lung cancer mortality rates among heavy smokers. Therefore, the need for low-dose CT screening should be considered internationally. In screening and routine clinical practice, the detection of numerous pulmonary nodules on chest radiography or chest CT is an important task for physicians. However, increased workload may lead to physician burnout [35].

Recent studies have shown that AI has high performance in detecting pulmonary nodules on chest radiographs and CT scans [36–41]. In the detection studies reviewed for chest radiography using the AI-based computer-aided detection/diagnosis (CAD) system [37, 39, 40], the sensitivity ranges from 79.0 to 91.1% and the specificity from 93 to 100%. The value of the false positive (FP) per scan is between 0.02 and 0.34. Moreover, the AI-based CAD system with bone suppression technology was more sensitive to detecting pulmonary nodules on chest radiographs compared with the original model by improving the visibility of pulmonary nodules hidden behind the clavicle and ribs (Fig. 2) [42]. On the other hand, in the detection studies reviewed for CT using the AI-based CAD system [37, 38, 43], the sensitivity ranges from 61.6 to 98.1%, and the value of the FP per scan ranges from 0.125 to 32. Pulmonary nodules have been



Fig. 2 Nodule detection on chest radiograph with bone suppression by a commercially available AI model. A 42-year-old woman. A low radiolucent area is suspected on the right supraclavicular bone (arrow, **A**). However, it is difficult to distinguish between pulmonary and bone lesions. A pulmonary nodule hidden behind the right clavicle

can be detected on the chest radiograph with bone suppression using a commercially available AI model (arrow, **B**). A part-solid ground-glass nodule with about 3 cm in diameter can be seen in the right upper lobe (arrow, **C**). This nodule was confirmed as a lung adenocarcinoma after surgery. *AI* artificial intelligence

already detected with high area under the curve (AUC) values close to 1. However, studies have reported a wide range of FP rates for AI detection of pulmonary nodules in chest radiographs and CT scans, as shown above [36–40]. The variability in reported FP rates may be due to differences in the AI models used, the populations studied, and the specific evaluation metrics used. It is important to note that reducing FP rates is a key challenge for improving the accuracy and effectiveness of AI-based nodule detection systems.

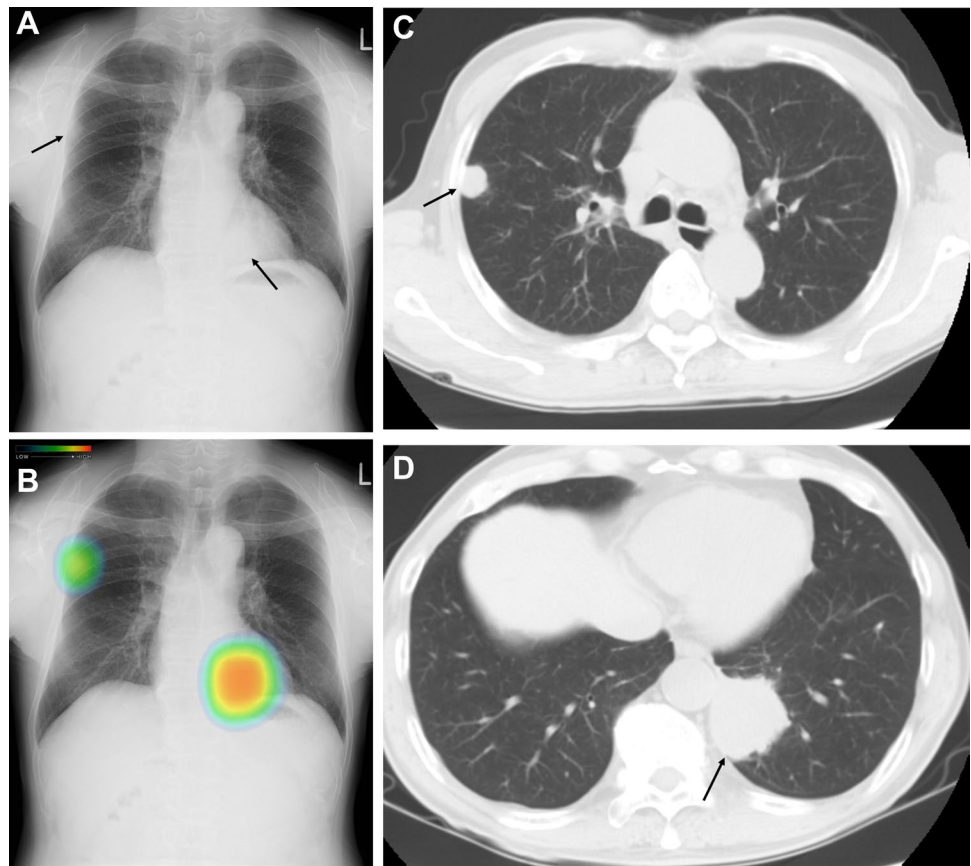
AI-based CAD system has already outperformed physicians in nodule detection performance on chest radiography and CT, and it enhanced physicians' performances when used as a second reader [40, 44]. The AI-based CAD system is useful in medical examinations and daily clinical practice to prevent physicians from overlooking pulmonary nodules and is also useful as a diagnostic support in areas with a shortage of physicians (Fig. 3). However, there are still many challenges and limitations remaining including overfitting, lack of interpretability, and insufficient annotated data. A generative model via an adversarial setting called generative adversarial network (GAN), proposed by Goodfellow et al. [45], has been widely used to generate synthetic images for medical model development. Madani et al. [46] demonstrated that semi-supervised GANs can achieve superior performance compared to conventional supervised CNNs, while using less training data. Although GANs are less prone to overfitting, their training process can be unstable due to the Nash equilibrium between the generator and the discriminator. The use of GANs in the medical domain may raise ethical concerns. In order to give AI the same sense of ethics as humans, it is crucial to improve the process by reviewing biases in the knowledge and information used for training in advance. Care must be taken to ensure that

seemingly positive results do not lead to biased outcomes and judgments. Further studies are needed to investigate the generalizability and clinical utility of an AI-based CAD system on chest radiography and CT.

Clinical application to lung tumors: diagnosability of pulmonary nodules

Detection of lung cancer in its early nodular stage through screening, categorization, and medical care has proven to be highly beneficial and effective in reducing mortality rates associated with lung cancer. Qualitative and quantitative evaluations using images play an important role in the diagnosis of thoracic tumors including lung cancer [47–61]. Diagnosis of the malignancy or benignity of pulmonary nodules is also one of the critical tasks for physicians, but it is commonly acknowledged that achieving 100% accuracy in determining the malignancy or benignity of nodules based on CT images alone is unattainable. As a result, in many cases, the diagnosis of CT images without histopathologic diagnosis requires reliance on follow-up observations. Therefore, it would be reassuring for physicians to have AI support for qualitative diagnosis. For the early detection and diagnosis of lung cancer, various AI-based CAD systems have been developed to classify benign and malignant lung nodules [62–65]: AI-based CAD system with 2-dimensional (2-D) or 3-D CNN network achieved good classification performance in malignant suspicion of pulmonary nodules with close to 90% classification accuracy. In past reports [62–65], most of the existing CNN-based approaches use a 2D-CNN network due to the advantages of low network complexity and fast computation. However, compared with the 2D-CNN network, the 3D-CNN network can encode

Fig. 3 Nodule detection on chest radiograph by a commercially available AI-based CAD system. A 70-year-old man. A nodule can be seen in the right middle lung field and a mass can be seen in the left lower lung field overlapping the cardiac shadow (arrows, **A**). A commercially available AI-based CAD system can highlight both the nodule and mass. A nodule with 2 cm in diameter can be seen in the right upper lobe on CT (arrow, **C**). A mass with 5 cm in diameter can be seen in the left lower lobe on CT (arrow, **D**). This mass was confirmed as a lung metastasis from renal cell carcinoma after biopsy. *AI* artificial intelligence, *CAD* computer-aided detection/diagnosis



richer spatial information and extract more representative features through its hierarchical architecture trained with 3D samples, which also contributes to the efficient detection of nodules to be qualitatively diagnosed by reducing the FP rates in automated pulmonary nodule detection on CT [66], while requiring a large amount of data and taking a long time to compute. Parallel processing of ViT models [28, 29] is expected to improve diagnostic performance in the future because it has an advantage in terms of computational cost and its performance does not deteriorate easily even if the model size is increased, but the need to prepare huge data for model training is the same as for CNN models. The potential benefits of ViT in improving diagnostic accuracy should be further investigated in the future.

As with nodule detection, an AI-based CAD system can have at least some impact on physician diagnosis. An AI model predicting the pathological invasiveness of lung adenocarcinoma, constructed from a small dataset of 285 cases, showed diagnostic performance with an accuracy of 74.2%, a sensitivity of 80.3%, a specificity of 67.1% [7], and an AUC value of 0.74. When this model was used by three different physicians with 9, 14, and 26 years of experience, sensitivity was improved at the expense of specificity. There was no significant difference in overall diagnostic performance, but the accuracy of the less experienced physician

was significantly improved. Although it is difficult for physicians to predict the invasiveness of lung adenocarcinoma manifested as ground-glass nodules on CT images, the combination of physicians with AI-based CAD system using CNN network can outperform physicians alone in internal dataset (AUC values, 0.819 vs. 0.606) and external dataset (AUC values, 0.893 vs. 0.693) [67]. In addition, CNN may be beneficial in improving the accuracy of computerized detection of volume measurements and the ability to differentiate nodules on CT scans in patients with pulmonary nodules [68]. Wataya et al. [69] demonstrated that a commercially available AI-based CAD system could measure the accuracy of imaging features of nodules/masses with high interobserver agreement, thereby improving the accuracy of benign or malignant diagnosis. In particular, the AI-based CAD system was valuable for radiologists with less experience. AI-based CAD system for diagnosis will be very promising in the future.

Clinical application to lung tumors: prediction of prognosis

Many approaches have been investigated to develop deep learning prognostic models for lung cancer patients using CT scans. In particular, the ability to predict preoperative

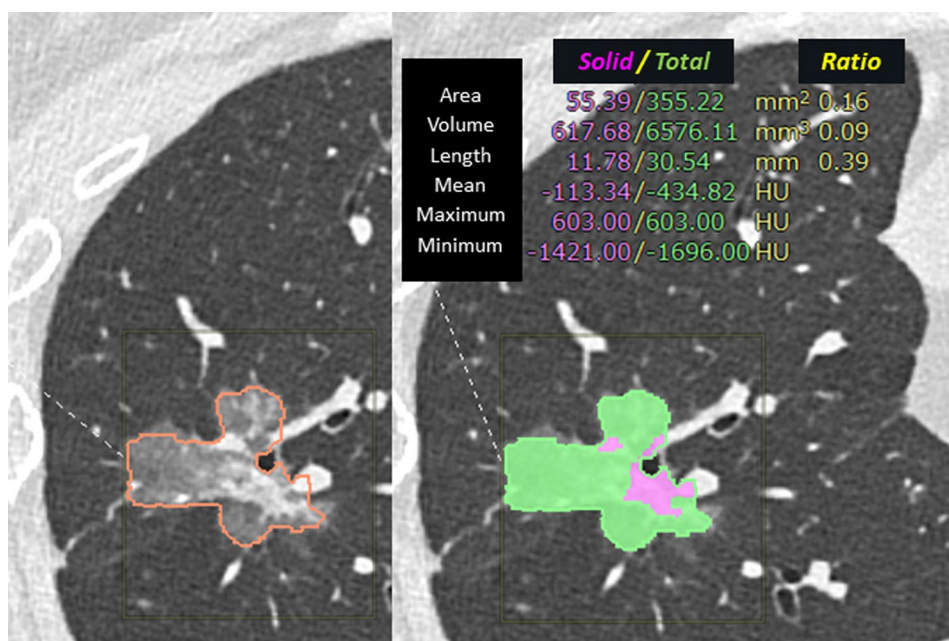
outcomes using CT data would be extremely beneficial for individuals diagnosed with early-stage lung cancer, as it would assist in determining appropriate treatment strategies [70–72].

In the clinical setting, measuring the size of invasiveness on CT for the T descriptor size is very important in the 8th edition of the TNM lung cancer classification because the T descriptor size correlates with the patients' prognosis [73, 74]. To decide the T descriptor of lung cancer, the solid component on CT is measured in centimeters (to one decimal place) with a slice thickness of 1 mm or less in the lung field conditions (recommended). Measurements are usually assessed in cross section, but the lack of standardized methods for tumor characterization and measurement remains a problem. Therefore, measurements are dependent on viewing conditions and observers, resulting in variability. Quantitative analysis using 3D volumetry is more advantageous in terms of objectivity and reproducibility of measurements [75, 76]. Wormanns et al. [76] demonstrated that a 25% increase in measured volume had a 95% probability of reflecting true nodule growth rather than measurement variability. However, differences in conditions such as software performance and nodule characteristics cause some errors in measurements, whether between the same observer or between different observers. Even if the software is the same, care must be taken when comparing measurements. On the other hand, a recent commercially available AI-based CAD system [69], which is already in clinical use in our country, can automatically provide quantitative values and exactly the same values regardless of display conditions and observers (Fig. 4). This AI-based CAD system can provide quantitative values with extremely high accuracy and

reproducibility. Volumetric measurements, such as solid volume and solid volume percentage, have been found to be independent indicators associated with an increased likelihood of recurrence and/or death in lung cancer patients [77]. Highly accurate quantification using an AI-based CAD system may lead to highly reproducible prognostic prediction in clinical practice.

A recent study put forth a deep learning model that directly predicted the cumulative survival probability in patients diagnosed with early stage lung adenocarcinoma using preoperative CT data [72]. Multivariate analysis data demonstrated smoking status (hazard ratio [HR], 3.4; $p=0.007$) and the output of the deep learning model (categorical form: HR, 3.6; $p=0.003$) were the only independent prognostic factors for disease-free survival in patients with clinical stage I lung adenocarcinoma. However, this approach had some limitations in terms such as the time-to-event data used to train the model and the presence of censoring, particularly in early-stage lung cancer patients. Prediction of prognosis directly from images may output highly reproducible results, but due to the nature of AI, whose analysis process is a black box, it may sometimes be difficult to accept the final results of AI. Zhong et al. [78] reported a prognostic stratification AI of lymph node metastasis using based on the AI score calculated from the CT image data of the primary lung cancer, which may be more acceptable to us. In the future, we hope to build a more accurate prognosis prediction model by learning clinical data such as many morphological evaluations and histological characteristics. In addition, so-called ablation studies, in which the elements that make up the model are deliberately deleted to evaluate how these elements contribute to the performance of the

Fig. 4 Quantification by a commercially available AI-based CAD system. A commercially available AI-based CAD system can automatically segment a part-solid ground-glass nodule, providing quantitative values that are exactly the same regardless of the observers. AI artificial intelligence, CAD computer-aided detection/diagnosis



model, are also considered important in the construction of the model.

Clinical application for infectious diseases: focusing on COVID-19

Although there are many respiratory infections, many infectious disease AIs naturally target COVID-19, formally known as severe acute respiratory syndrome coronavirus 2 [79, 80]. In general, chest radiography and CT are indispensable for diagnostic imaging of chest diseases including infectious diseases [81, 82]. The usefulness of chest radiography and CT for COVID-19 associated with diagnosis, disease severity, and prognosis has been reported [83–90]. Radiologists exhibit a remarkably high proficiency in identifying COVID-19 pneumonia through the assessment of chest CT scans [91]. To combat the rapid spread of COVID-19, many AI models using imaging findings have been developed worldwide [12, 80]. Ito et al. demonstrated that evaluating the use of AI in diagnostic imaging for COVID-19 patients has the potential to serve as a pilot study for determining the ideal AI systems to be employed during the initial phases of future emerging diseases, providing valuable insights into the requirements for dataset size and the appropriate application of AI in such scenarios [80].

As an initial representative paper using chest CT, Li et al. [92] created an AI program using CT scans to identify individuals with COVID-19 pneumonia. The program demonstrated a sensitivity of 90%, a specificity of 96%, and an AUC of 0.96 for diagnosing COVID-19 pneumonia. A dataset of 400 COVID-19 patients, 1396 patients with community-acquired pneumonia, and 1173 patients with normal CT scans or without pneumonia was used to train the program. The performance of the program was assessed using a separate dataset consisting of 68 COVID-19 patients, 155 patients with community-acquired pneumonia, and 130 patients with normal CT scans or without pneumonia. The COVID-19 cases in the dataset were confirmed through reverse transcription polymerase chain reaction test. Although the specific details of the dataset were not disclosed, the source code was made available and could be visualized using gradient-weighted class activation mapping (Grad-CAM). The study encompassed a large dataset and achieved a high level of accuracy. However, various AI models for COVID-19 pneumonia classification using chest radiography and chest CT that have appeared since the early days of the COVID-19 have been in a mixed state, with conspicuous variations in databases and analysis methods [80, 93]. Early-stage AI results may have been more influenced by morbidity than by imaging findings. Hence, both sensitivity and specificity ranged from high to low compared to other AI creation researches, which seems to indicate the difficulty of actual AI operation. On the other hand, recent

AI models of COVID-19 pneumonia have improved the accuracy of diagnosis and severity prediction through the use of highly accurate lesion segmentation techniques, quantitative methods including radiomics approaches, and combination with clinical information. The meta-analysis [94] revealed that AI models based on chest imaging discriminate COVID-19 from other pneumonias: pooled AUC of 0.96 (95% confidence interval [CI] 0.94–0.98), pooled sensitivity of 0.92 (95% CI 0.88–0.94), pooled specificity of 0.91 (95% CI 0.87–0.93). As in other areas, future researchers should pay more attention to the quality of research methodology and further improve the generalizability of the developed predictive models, while considering the importance of database construction, uniformity of evaluation methods, and the necessity of benchmarking the evaluation database.

Clinical application for non-tumor/non-infection: focusing on pulmonary embolism

This section focused on PE, although there are many reports on the usefulness of thoracic imaging in non-oncologic and non-infectious diseases [95–97]. PE carries a high burden of morbidity and mortality. There is also a report demonstrated that the detection of an incidental PE on a staging CT scan is associated with a very high risk of progressive malignancy [98]. Timely and precise diagnosis enables the acceleration of treatment, a crucial factor in potentially reducing mortality rates and enhancing patient outcomes [99–101]. CT pulmonary angiography (CTPA) is the diagnostic standard for noninvasive confirmation of PE. Either dual-energy or single-energy CT can be used comparably for the detection of acute PE [102, 103]. Nevertheless, the identification of PE through CTPA is a tedious procedure that necessitates the expertise of radiologists. Consequently, the interpretation process is prone to error and may lead to delayed diagnoses.

An innovative approach to PE detection has been developed using AI, which includes three main tasks: classification to label an entire image, detection to localize an individual object in the image, and segmentation to delineate a pixel-wise border of emboli. Meta-analysis results [104] demonstrated that AI-based CAD system for PE indicated a pooled sensitivity of 0.88 (95% CI 0.803–0.927) per scan and a specificity of 0.86 (95% CI 0.756–0.924) per scan, which showed a better balance between sensitivity and specificity compared to radiologists' sensitivity of 0.67–0.87 with a specificity of 0.89–0.99. The effective AI models will support physicians by reducing the rate of missed findings, thereby helping to minimize the time required to review the CT scans. Cheikh et al. [105] demonstrated that AI for the detection of PE seemed to serve as a safety net rather than a replacement for radiologists in emergency radiology practice. This is attributed to its high sensitivity and negative predictive value (NPV), which subsequently increases the

confidence of radiologists. Batra et al. [106] retrospectively evaluated the performance of an AI algorithm in detecting incidental PE on enhanced chest CT examinations in 2555 patients. The AI algorithm was applied to the images, and a natural language processing (NLP) algorithm was used to analyze the clinical reports. A multi-reader adjudication process was implemented to establish a reference standard for incidental PE. The AI detected 4 incidental PEs missed by clinical reports, while the clinical reports identified 7 incidental PEs missed by the AI. The AI exhibited lower positive predictive value (PPV) and specificity than clinical reports, but there were no significant differences in sensitivity and NPV. The AI showed high NPV and moderate PPV in detecting incidental PE, complementing the radiologists' findings by identifying some cases missed by them. Potential applications of the AI model serve as a second reader to help allow earlier PE detection and intervention, and improving the model by understanding AI misclassifications.

Patients with COVID-19 pneumonia are susceptible to developing deep vein thrombosis and PE. Prophylactic use of low-molecular-weight heparin does not reduce the risk of venous thromboembolism in COVID-19 pneumonia [107, 108]. Therefore, the performance of a commercially available AI algorithm for PE detection on enhanced CT in patients hospitalized for COVID-19 was investigated [109]. The AI algorithm demonstrated high sensitivity of 93.2% (95% CI 0.906–0.952) and specificity of 99.6% (95% CI 0.989–0.999) for PE on enhanced CT scans in patients with COVID-19 regardless of parenchymal disease. However, accuracy was significantly affected by the mean attenuation of the pulmonary vasculature, which may warrant further investigation in the future.

Notes and expectations for future AI

The use of AI techniques in thoracic CT imaging has been shown to be beneficial for a number of fronts such as detection, diagnosis, and prognosis. AI can provide this information in a more objective, reproducible, and robust manner compared to subjective analysis and conventional quantification. Overfitting as one of the problems of AI, which is learned on the training data but not adapted to the unknown data, still remains. Building AI models by training on a large amount of data is ideal, but it is difficult to create a perfect one. Unsupervised learning is used when labeled data is limited or when the goal is to understand the features and structures of the data or to gain new insights. However, compared to supervised learning, unsupervised learning poses challenges in evaluating predictive performance and optimization. On the other hand, simply adding more training data does not always seem appropriate. CNNs trained on a modest collection of

prospectively labeled data might achieve high diagnostic performance in classifying their data as normal or abnormal [110]. Care should also be taken when enriching data to improve generalizability. Methodological pitfalls have the potential to produce a superficially well-performed AI model because oversampling or augmenting data before it is divided into training, test, and validation data violates data independence [111].

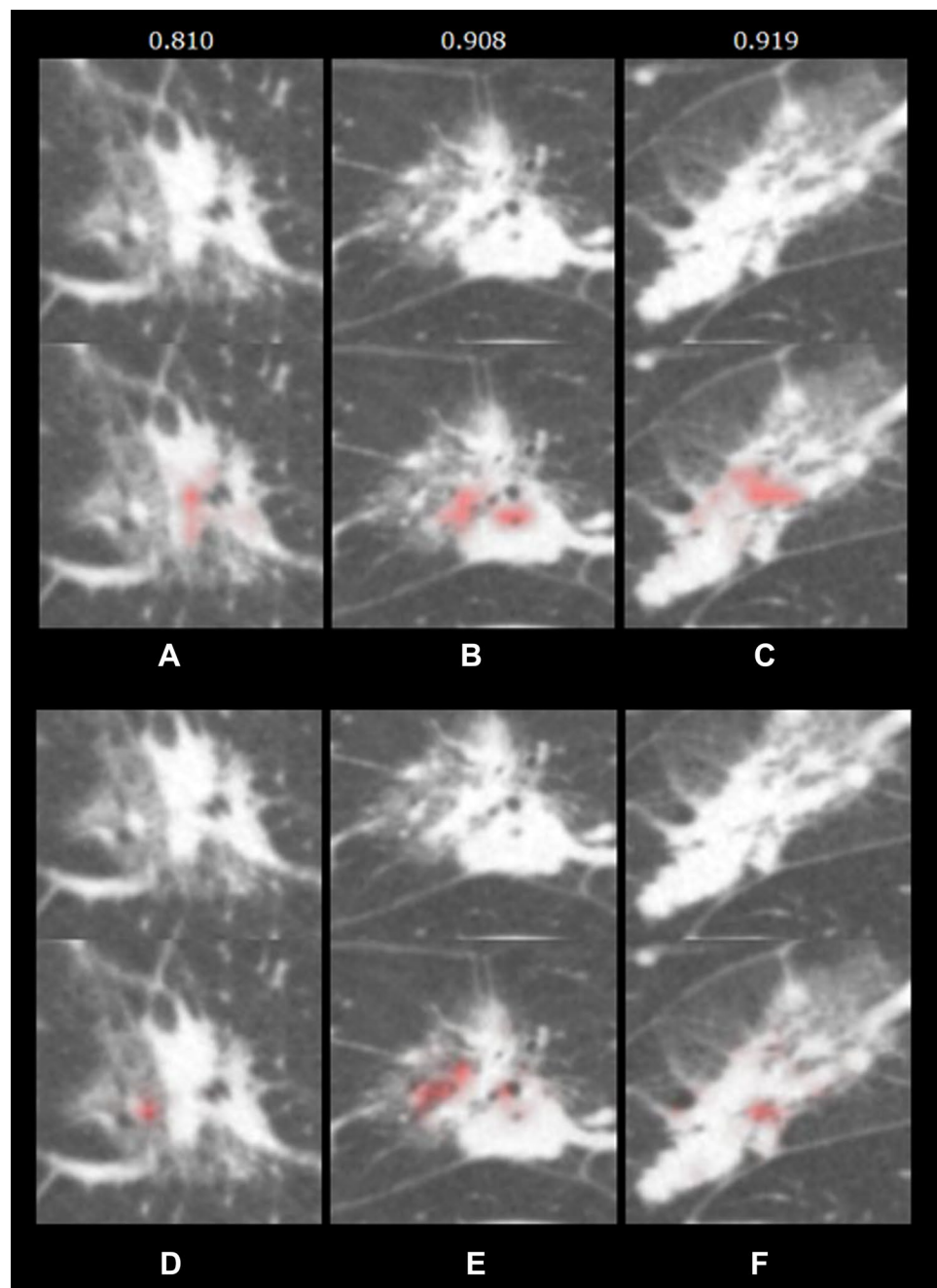
Another AI problem is the black-boxing of the analysis process.

It is possible to identify features within the image on which DL focuses by using Grad-CAMs, attention map, and local interpretable model-agnostic explanations (LIME).

These heat maps cannot tell us what an AI concretely analyzed, but at least tell us where it focused (Fig. 5), which might prevent the possibility of unexpected findings being used to make a final decision due to the black box in the process of an AI. Further quantitative analysis of the focused lesions and correlation with pathology may be useful in constructing an explainable AI. It is desirable to build an end-to-end explainable AI that integrates various functions.

Imaging in the thoracic field requires a wide range of roles, not only morphological diagnosis but also functional analysis and prognosis prediction. There is a greater potential for AI to contribute to thoracic imaging. In addition, the development and validation of further AI technologies may be possible with new imaging techniques such as photon-counting CT [112, 113]. An AI chatbot based on a generative pretrained transformer model (ChatGPT), created by Open AI in 2018, can provide answers across a broad range of diverse topics. While the occurrence of generating inaccurate responses is reduced in a new version GPT-4, it still hampers the practicality of utilizing it in medical education and practice currently [114, 115]. In the future, the combination of GPT-4 and an AI-based CAD system may possibly provide structure reports, which may be useful in both clinical and research support. Furthermore, unlike the traditional AI, Hippocratic AI is receiving a lot of attention as one of the large language models with expertise in tasks related to medical information, such as patient diagnosis, treatment, and answering medical questions. Hippocratic AI is designed to protect patient data and take security measures to ensure proper handling of patients' personal information, which might provide accurate information and up-to-date medical knowledge for valuable support to healthcare professionals and patients [116]. AI is expected to enable more accurate medical care and provide better personalized medical services to patients. Last of all, it is crucial to remember that AI functions as a tool to assist physicians, and the ultimate decision should be made by the physician while comprehending the limitations of AI.

Fig. 5 Differences of attention map in the diagnosis of lung nodule. This is a case confirmed as invasive adenocarcinoma after surgery (**a** and **d**, axial images; **b** and **e**, coronal images; **c** and **f**, sagittal images). Red areas in **a**, **b**, and **c** are focused by the AI in the diagnosis of invasive adenocarcinoma. Even for the same nodule, red areas in **d**, **e**, and **f** are focused when diagnosed as noninvasive adenocarcinoma. The numbers in the upper row indicate the diagnostic probability of invasive adenocarcinoma (maximum value of 1) calculated by the AI. Depending on the purpose of diagnosis and the cross section of the image, the focused part and the probability of results by the AI may differ. *AI* artificial intelligence



Authors' contributions The first draft of the manuscript was written by MY and all authors commented on previous versions of the manuscript. All authors read and approved the final version of the manuscript.

Funding Open access funding provided by Osaka University. This review article has not received any specific grant from any public, commercial, or nonprofit funding agency.

Declarations

Conflict of interest All authors declare no conflicts of interest about this manuscript.

Ethical standards This article does not contain any studies with human participants or animals performed by any of the authors.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing,

adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Crevier D (1993) AI: the tumultuous search for artificial intelligence. BasicBooks, New York, NY
2. McCorduck, Pamela (2004) *Machines Who Think* (2nd ed.), Natick, MA: A. K. Peters, Ltd., ISBN 1-56881-205-1, OCLC 52197627
3. Russell S, Norvig P (2020) *Artificial intelligence: a modern approach*, 4th edn. Pearson, London, pp 19–53
4. Dhawan AP (2011) *Medical image analysis*. Wiley, Hoboken, NJ
5. Suzuki K (2017) Overview of deep learning in medical imaging. *Radiol Phys Technol* 10(3):257–273. <https://doi.org/10.1007/s12194-017-0406-5>
6. Yanagawa M, Niioka H, Hata A, Kikuchi N, Honda O, Kurakami H, Morii E, Noguchi M, Watanabe Y, Miyake J, Tomiyama N (2019) Application of deep learning (3-dimensional convolutional neural network) for the prediction of pathological invasiveness in lung adenocarcinoma: a preliminary study. *Medicine (Baltimore)* 98(25):e16119. <https://doi.org/10.1097/MD.00000000000016119>
7. Yanagawa M, Niioka H, Kusumoto M, Awai K, Tsubamoto M, Satoh Y, Miyata T, Yoshida Y, Kikuchi N, Hata A, Yamasaki S, Kido S, Nagahara H, Miyake J, Tomiyama N (2021) Diagnostic performance for pulmonary adenocarcinoma on CT: comparison of radiologists with and without three-dimensional convolutional neural network. *Eur Radiol* 31:1978–1986. <https://doi.org/10.1007/s00330-020-07339-x>
8. Barat M, Chassagnon G, Dohan A, Gaujoux S, Coriat R, Hoeffel C, Cassinotto C, Soyer P (2021) Artificial intelligence: a critical review of current applications in pancreatic imaging. *Jpn J Radiol* 39:514–523. <https://doi.org/10.1007/s11604-021-01098-5>
9. Wong LM, Ai QYH, Mo FKF, Poon DMC, King AD (2021) Convolutional neural network in nasopharyngeal carcinoma: how good is automatic delineation for primary tumor on a non-contrast-enhanced fat-suppressed T2-weighted MRI? *Jpn J Radiol* 39:571–579. <https://doi.org/10.1007/s11604-021-01092-x>
10. Ichikawa Y, Kanii Y, Yamazaki A, Nagasawa N, Nagata M, Ishida M, Kitagawa K, Sakuma H (2021) Deep learning image reconstruction for improvement of image quality of abdominal computed tomography: comparison with hybrid iterative reconstruction. *Jpn J Radiol* 39:598–604. <https://doi.org/10.1007/s11604-021-01089-6>
11. Nakai H, Fujimoto K, Yamashita R, Sato T, Someya Y, Taura K, Isoda H, Nakamoto Y (2021) Convolutional neural network for classifying primary liver cancer based on triple-phase CT and tumor marker information: a pilot study. *Jpn J Radiol* 39:690–702. <https://doi.org/10.1007/s11604-021-01106-8>
12. Okuma T, Hamamoto S, Maebayashi T, Taniguchi A, Hirakawa K, Matsushita S, Matsushita K, Murata K, Manabe T, Miki Y (2021) Quantitative evaluation of COVID-19 pneumonia severity by CT pneumonia analysis algorithm using deep learning technology and blood test results. *Jpn J Radiol* 39:956–965. <https://doi.org/10.1007/s11604-021-01134-4>
13. Özer H, Kılınçer A, Uysal E, Yormaz B, Cebeci H, Durmaz MS, Koplay M (2021) Diagnostic performance of Radiological Society of North America structured reporting language for chest computed tomography findings in patients with COVID-19. *Jpn J Radiol* 39:877–888. <https://doi.org/10.1007/s11604-021-01128-2>
14. Kitahara H, Nagatani Y, Otani H, Nakayama R, Kida Y, Sonoda A, Watanabe Y (2022) A novel strategy to develop deep learning for image super-resolution using original ultra-high-resolution computed tomography images of lung as training dataset. *Jpn J Radiol* 40:38–47. <https://doi.org/10.1007/s11604-021-01184-8>
15. Yasaka K, Akai H, Sugawara H, Tajima T, Akahane M, Yoshioka N, Kabasawa H, Miyo R, Ohtomo K, Abe O, Kiryu S (2022) Impact of deep learning reconstruction on intracranial 1.5 T magnetic resonance angiography. *Jpn J Radiol* 40:476–483. <https://doi.org/10.1007/s11604-021-01225-2>
16. Kaga T, Noda Y, Mori T, Kawai N, Miyoshi T, Hyodo F, Kato H, Matsuo M (2022) Unenhanced abdominal low-dose CT reconstructed with deep learning-based image reconstruction: image quality and anatomical structure depiction. *Jpn J Radiol* 40:703–711. <https://doi.org/10.1007/s11604-022-01259-0>
17. Nakao T, Hanaoka S, Nomura Y, Hayashi N, Abe O (2022) Anomaly detection in chest 18F-FDG PET/CT by Bayesian deep learning. *Jpn J Radiol* 40:730–739. <https://doi.org/10.1007/s11604-022-01249-2>
18. Ohno Y, Aoyagi K, Arakita K, Doi Y, Kondo M, Banno S, Kasahara K, Ogawa T, Kato H, Hase R, Kashizaki F, Nishi K, Kamio T, Mitamura K, Ikeda N, Nakagawa A, Fujisawa Y, Taniguchi A, Ikeda H, Hattori H, Murayama K, Toyama H (2022) Newly developed artificial intelligence algorithm for COVID-19 pneumonia: utility of quantitative CT texture analysis for prediction of favipiravir treatment effect. *Jpn J Radiol* 40:800–813. <https://doi.org/10.1007/s11604-022-01270-5>
19. Ozaki J, Fujioka T, Yamaga E, Hayashi A, Kujiraoka Y, Imokawa T, Takahashi K, Okawa S, Yashima Y, Mori M, Kubota K, Oda G, Nakagawa T, Tateishi U (2022) Deep learning method with a convolutional neural network for image classification of normal and metastatic axillary lymph nodes on breast ultrasonography. *Jpn J Radiol* 40:814–822. <https://doi.org/10.1007/s11604-022-01261-6>
20. Cay N, Mendi BAR, Batur H, Erdogan F (2022) Discrimination of lipoma from atypical lipomatous tumor/well-differentiated liposarcoma using magnetic resonance imaging radiomics combined with machine learning. *Jpn J Radiol* 40:951–960. <https://doi.org/10.1007/s11604-022-01278-x>
21. Han D, Chen Y, Li X, Li W, Zhang X, He T, Yu Y, Dou Y, Duan H, Yu N (2023) Development and validation of a 3D-convolutional neural network model based on chest CT for differentiating active pulmonary tuberculosis from community-acquired pneumonia. *Radiol Med* 128:68–80. <https://doi.org/10.1007/s11547-022-01580-8>
22. Abdullah SS, Rajasekaran MP (2022) Automatic detection and classification of knee osteoarthritis using deep learning approach. *Radiol Med* 127:398–406. <https://doi.org/10.1007/s11547-022-01476-7>
23. Nai YH, Loi HY, O'Doherty S, Tan TH, Reilhac A (2022) Comparison of the performances of machine learning and deep learning in improving the quality of low dose lung cancer PET images. *Jpn J Radiol* 40(12):1290–1299. <https://doi.org/10.1007/s11604-022-01311-z>
24. Bai HX, Wang R, Xiong Z, Hsieh B, Chang K, Halsey K, Tran TML, Choi JW, Wang DC, Shi LB, Mei J, Jiang XL, Pan I, Zeng QH, Hu PF, Li YH, Fu FX, Huang RY, Sebros R, Yu QZ, Atalay MK, Liao WH (2020) Artificial intelligence augmentation of

- radiologist performance in distinguishing COVID-19 from pneumonia of other origin at chest CT. *Radiology* 296:E156–E165. <https://doi.org/10.1148/radiol.2020201491>
25. Park SH, Han K (2018) Methodologic guide for evaluating clinical performance and effect of artificial intelligence technology for medical diagnosis and prediction. *Radiology* 286(3):800–809. <https://doi.org/10.1148/radiol.2017171920>
 26. Zech JR, Badgeley MA, Liu M, Costa AB, Titano JJ, Oermann EK (2018) Variable generalization performance of a deep learning model to detect pneumonia in chest radiographs: a cross-sectional study. *PLoS Med.* 15(11):e1002683
 27. Oakden-Rayner L, Dunnmon J, Carneiro G, Ré C (2020) Hidden stratification causes clinically meaningful failures in machine learning for medical imaging. *Proc ACM Conf Health Inference Learn* 2020:151–159. <https://doi.org/10.1145/3368555.3384468>
 28. Dosovitskiy A, Beyer L, Kolesnikov A, Weissenborn D, Zhai X, Unterthiner T, Dehghani M, Minderer M, Heigold G, Gelly S, Uszkoreit J, Houlsby N (2022) An Image is worth 16 × 16 words: transformers for image recognition at scale. <https://arxiv.org/abs/2010.11929> (preprint)
 29. Touvron H, Cord M, Douze M, Massa F, Sablayrolles A, Jegou H (2022). Training data-efficient image transformers & distillation through attention. <https://arxiv.org/abs/2012.12877> (preprint)
 30. Murphy ZR, Venkatesh K, Sulam J, Yi PH (2022) Visual transformers and convolutional neural networks for disease classification on radiographs: a comparison of performance, sample efficiency, and hidden stratification. *Radiol Artif Intell* 4:e220012. <https://doi.org/10.1148/ryai.220012>
 31. Rajaraman S, Zamzmi G, Folio LR, Antani S (2022) Detecting Tuberculosis-consistent findings in lateral chest X-rays using an ensemble of CNNs and vision transformers. *Front Genet* 13:864724. <https://doi.org/10.3389/fgene.2022.864724>
 32. Siegel RL, Miller KD, Fuchs HE, Jemal A (2022) Cancer statistics, 2022. *CA Cancer J Clin* 72(1):7–33. <https://doi.org/10.3322/caac.21708>
 33. National Lung Screening Trial Research Team, Aberle DR, Adams AM et al (2011) Reduced lung-cancer mortality with low-dose computed tomographic screening. *N Engl J Med* 365:395–409. <https://doi.org/10.1056/NEJMoa1102873>
 34. de Koning HJ, van der Aalst CM, de Jong PA et al (2020) Reduced lung-cancer mortality with volume CT screening in a randomized trial. *N Engl J Med* 382:503–513. <https://doi.org/10.1056/NEJMoa1911793>
 35. Parikh JR, Wolfman D, Bender CE, Arleo E (2020) Radiologist burnout according to surveyed radiology practice leaders. *J Am Coll Radiol* 17(1 Pt A):78–81. <https://doi.org/10.1016/j.jacr.2019.07.008>
 36. Agrawal T, Choudhary P (2023) Segmentation and classification on chest radiography: a systematic survey. *Vis Comput* 39(3):875–913. <https://doi.org/10.1007/s00371-021-02352-7>
 37. Sun Y, Li C, Zhang Q, Zhou A, Zhang G (2021). Survey of the detection and classification of pulmonary lesions via CT and X-ray. <https://doi.org/10.48550/arXiv.2012.15442>
 38. Gu Y, Chi J, Liu J, Yang L, Zhang B, Yu D, Zhao Y, Lu X (2021) A survey of computer-aided diagnosis of lung nodules from CT scans using deep learning. *Comput Biol Med* 137:104806. <https://doi.org/10.1016/j.compbimed.2021.104806>
 39. Li X, Shen L, Xie X et al (2019) Multi-resolution convolutional networks for chest X-ray radiograph based lung nodule detection. *Artif Intell Med* 103:101744. <https://doi.org/10.1016/j.artmed.2019.101744>
 40. Nam JG, Park S, Hwang EJ et al (2019) Development and validation of deep learning-based automatic detection algorithm for malignant pulmonary nodules on chest radiographs. *Radiology* 290:218–228. <https://doi.org/10.1148/radiol.2018180237>
 41. Toda N, Hashimoto M, Iwabuchi Y, Nagasaka M, Takeshita R, Yamada M, Yamada Y, Jinzaki M (2023) Validation of deep learning-based computer-aided detection software use for interpretation of pulmonary abnormalities on chest radiographs and examination of factors that influence readers' performance and final diagnosis. *Jpn J Radiol* 41(1):38–44. <https://doi.org/10.1007/s11604-022-01330-w>
 42. Kim H, Lee KH, Han K et al (2023) Development and validation of a deep learning-based synthetic bone-suppressed model for pulmonary nodule detection in chest radiographs. *JAMA Netw Open* 6(1):e2253820. <https://doi.org/10.1001/jamanetworkopen.2022.53820>
 43. Takaishi T, Ozawa Y, Bando Y, Yamamoto A, Okochi S, Suzuki H, Shibamoto Y (2021) Incorporation of a computer-aided vessel-suppression system to detect lung nodules in CT images: effect on sensitivity and reading time in routine clinical settings. *Jpn J Radiol* 39(2):159–164. <https://doi.org/10.1007/s11604-020-01043-y>
 44. Katase S, Ichinose A, Hayashi M et al (2022) Development and performance evaluation of a deep learning lung nodule detection system. *BMC Med Imaging* 22(1):203. <https://doi.org/10.1186/s12880-022-00938-8>
 45. Goodfellow IJ, Pouget-Abadie J, Mirza M, Xu B, WardeFarley D, Ozair S, Courville A, Bengio Y (2014) Generative adversarial nets. In: *Proceedings of the 27th international conference on neural information processing systems*, vol. 2, pp. 2672–2680
 46. Madani A, Moradi M, Karargyris A, Syeda-Mahmood T (2018) Semi-supervised learning with generative adversarial networks for chest x-ray classification with ability of data domain adaptation. In: *2018 IEEE 15th international symposium on biomedical imaging*. ISBI, pp. 1038–1042
 47. Suzuki K, Koike T, Asakawa T et al (2011) A prospective radiological study of thin-section computed tomography to predict pathological noninvasiveness in peripheral clinical IA lung cancer (Japan Clinical Oncology Group 0201). *J Thorac Oncol* 6(4):751–756. <https://doi.org/10.1097/JTO.0b013e31821038ab>
 48. Aoki T, Hanamiya M, Uramoto H, Hisaoka M, Yamashita Y, Korogi Y (2012) Adenocarcinomas with predominant ground-glass opacity: correlation of morphology and molecular biomarkers. *Radiology* 264(2):590–596. <https://doi.org/10.1148/radiol.12111337>
 49. Aokage K, Yoshida J, Ishii G, Matsumura Y, Haruki T, Hishida T et al (2013) Identification of early t1b lung adenocarcinoma based on thin-section computed tomography findings. *J Thorac Oncol* 8(10):1289–1294
 50. Yanagawa M, Tsubamoto M, Satoh Y, Hata A, Miyata T, Yoshida Y, Kikuchi N, Kurakami H, Tomiyama N (2020) Lung adenocarcinoma at CT with 0.25-mm section thickness and a 2048 matrix: high-spatial-resolution imaging for predicting invasiveness. *Radiology* 297(2):462–471. <https://doi.org/10.1148/radiol.2020201911>
 51. Kim SK, Kim TJ, Chung MJ, Kim TS, Lee KS, Zo JI, Shim YM (2018) Lung adenocarcinoma: CT features associated with spread through air spaces. *Radiology* 289(3):831–840. <https://doi.org/10.1148/radiol.2018180431>
 52. Jiang C, Luo Y, Yuan J, You S, Chen Z, Wu M, Wang G, Gong J (2020) CT-based radiomics and machine learning to predict spread through air space in lung adenocarcinoma. *Eur Radiol* 30(7):4050–4057. <https://doi.org/10.1007/s00330-020-06694-z>
 53. Chen Y, Jiang C, Kang W, Gong J, Luo D, You S, Cheng Z, Luo Y, Wu K (2022) Development and validation of a CT-based nomogram to predict spread through air space (STAS) in peripheral stage IA lung adenocarcinoma. *Jpn J Radiol* 40(6):586–594. <https://doi.org/10.1007/s11604-021-01240-3>
 54. Sadohara J, Fujimoto K, Müller NL, Kato S, Takamori S, Ohkuma K, Terasaki H, Hayabuchi N (2006) Thymic epithelial

- tumors: comparison of CT and MR imaging findings of low-risk thymomas, high-risk thymomas, and thymic carcinomas. *Eur J Radiol* 60(1):70–79. <https://doi.org/10.1016/j.ejrad.2006.05.003>
55. Ozawa Y, Hiroshima M, Maki H, Hara M, Shibamoto Y (2021) Imaging findings of lesions in the middle and posterior mediastinum. *Jpn J Radiol* 39(1):15–31. <https://doi.org/10.1007/s11604-020-01025-0>
 56. Mogami H, Onoike Y, Miyano H, Arakawa K, Inoue H, Sakae K, Kawakami T (2021) Lung cancer screening by single-shot dual-energy subtraction using flat-panel detector. *Jpn J Radiol* 39(12):1168–1173. <https://doi.org/10.1007/s11604-021-01163-z>
 57. Xu L, Lin S, Zhang Y (2022) Differentiation of adenocarcinoma in situ with alveolar collapse from minimally invasive adenocarcinoma or invasive adenocarcinoma appearing as part-solid ground-glass nodules (≤ 2 cm) using computed tomography. *Jpn J Radiol* 40(1):29–37. <https://doi.org/10.1007/s11604-021-01183-9>
 58. Jiang J, Fu Y, Zhang L, Liu J, Gu X, Shao W, Cui L, Xu G (2022) Volumetric analysis of intravoxel incoherent motion diffusion-weighted imaging in preoperative assessment of non-small cell lung cancer. *Jpn J Radiol* 40(9):903–913. <https://doi.org/10.1007/s11604-022-01279-w>
 59. Zhu Y, Lv W, Wu H, Yang D, Nie F (2022) A preoperative nomogram for predicting the risk of sentinel lymph node metastasis in patients with T1-2N0 breast cancer. *Jpn J Radiol* 40(6):595–606. <https://doi.org/10.1007/s11604-021-01236-z>
 60. Yoshifuji K, Toya T, Yanagawa N, Sakai F, Nagata A, Sekiya N, Ohashi K, Doki N (2021) CT classification of acute myeloid leukemia with pulmonary infiltration. *Jpn J Radiol* 39(11):1049–1058. <https://doi.org/10.1007/s11604-021-01151-3>
 61. Lin LY, Zhang F, Yu Y, Fu YC, Tang DQ, Cheng JJ, Wu HW (2022) Noninvasive evaluation of hypoxia in rabbit VX2 lung transplant tumors using spectral CT parameters and texture analysis. *Jpn J Radiol* 40(3):289–297
 62. Lv E, Liu W, Wen P, Kang X (2021) Classification of benign and malignant lung nodules based on deep convolutional network feature extraction. *J Healthc Eng* 2021:8769652. <https://doi.org/10.1155/2021/8769652>
 63. Nibali A, He Z, Wollersheim D (2017) Pulmonary nodule classification with deep residual networks. *Int J Comput Assist Radiol Surg* 12(10):1799–1808. <https://doi.org/10.1007/s11548-017-1605-6>
 64. Tran GS, Nghiem TP, Nguyen VT, Luong C, Burie J (2019) Improving accuracy of lung nodule classification using deep learning with focal loss. *J Healthc Eng*. <https://doi.org/10.1155/2019/5156416>
 65. Zhao J, Zhang C, Li D, Niu J (2020) Combining multi-scale feature fusion with multi-attribute grading, a CNN model for benign and malignant classification of pulmonary nodules. *J Digit Imaging* 33(4):869–878. <https://doi.org/10.1007/s10278-020-00333-1>
 66. Dou Q, Chen H, Yu L, Qin J, Heng PA (2017) Multilevel contextual 3-D CNNs for false positive reduction in pulmonary nodule detection. *IEEE Trans Biomed Eng* 64(7):1558–1567. <https://doi.org/10.1109/TBME.2016.2613502>
 67. Zhang T, Wang Y, Sun Y, Yuan M, Zhong Y, Li H, Yu T, Wang J (2021) High-resolution CT image analysis based on 3D convolutional neural network can enhance the classification performance of radiologists in classifying pulmonary non-solid nodules. *Eur J Radiol* 141:109810. <https://doi.org/10.1016/j.ejrad.2021>
 68. Ohno Y, Aoyagi K, Yaguchi A, Seki S, Ueno Y, Kishida Y, Takenaka D, Yoshikawa T (2020) Differentiation of benign from malignant pulmonary nodules by using a convolutional neural network to determine volume change at chest CT. *Radiology* 296(2):432–443. <https://doi.org/10.1148/radiol.2020191740>
 69. Wataya T, Yanagawa M, Tsubamoto M, Sato T, Nishigaki D, Kita K, Yamagata K, Suzuki Y, Hata A, Kido S, Tomiyama N (2023) Radiologists with and without deep learning-based computer-aided diagnosis: comparison of performance and interobserver agreement for characterizing and diagnosing pulmonary nodules/masses. *Eur Radiol* 33(1):348–359. <https://doi.org/10.1007/s00330-022-08948-4>
 70. Hosny A, Parmar C, Coroller TP, Grossmann P, Zeleznik R, Kumar A, Bussink J, Gillies RJ, Mak RH, Aerts HJWL (2018) Deep learning for lung cancer prognostication: a retrospective multi-cohort radiomics study. *PLoS Med* 15(11):e1002711. <https://doi.org/10.1371/journal.pmed.1002711>
 71. Torres FS, Akbar S, Raman S, Yasufuku K, Schmidt C, Hosny A, Baldauf-Lenschen F, Leighl NB (2021) End-to-end non-small-cell lung cancer prognostication using deep learning applied to pretreatment computed tomography. *JCO Clin Cancer Inform* 5:1141–1150. <https://doi.org/10.1200/CCI.21.00096>
 72. Kim H, Goo JM, Lee KH, Kim YT, Park CM (2020) Preoperative CT-based deep learning model for predicting disease-free survival in patients with lung adenocarcinomas. *Radiology* 296(1):216–224. <https://doi.org/10.1148/radiol.2020192764>
 73. Travis WD, Asamura H, Bankier AA, Beasley MB, Dettnerbeck F, Flieder DB, Goo JM, MacMahon H, Naidich D, Nicholson AG, Powell CA, Prokop M, Rami-Porta R, Rusch V, van Schil P, Yatabe Y, International Association for the Study of Lung Cancer Staging and Prognostic Factors Committee and Advisory Board Members (2016) The IASLC lung cancer staging project: proposals for coding T categories for subsolid nodules and assessment of tumor size in part-solid tumors in the forthcoming eighth edition of the TNM classification of lung cancer. *J Thorac Oncol* 11(8):1204–1223. <https://doi.org/10.1016/j.jtho.2016.03.025>
 74. Dettnerbeck FC, Bolejack V, Arenberg DA, Crowley J, Donington JS, Franklin WA, Girard N, Marom EM, Mazzzone PJ, Nicholson AG, Rusch VW, Tanoue LT, Travis WD, Asamura H, Rami-Porta R, IASLC Staging and Prognostic Factors Committee; Advisory Boards; Multiple Pulmonary Sites Workgroup; Participating Institutions (2016) The IASLC lung cancer staging project: background data and proposals for the classification of lung cancer with separate tumor nodules in the forthcoming eighth edition of the TNM classification for lung cancer. *J Thorac Oncol* 11(5):681–692. <https://doi.org/10.1016/j.jtho.2015.12.114>
 75. Yankelevitz DF, Reeves AP, Kostis WJ, Zhao B, Henschke CI (2000) Small pulmonary nodules: volumetrically determined growth rates based on CT evaluation. *Radiology* 217(1):251–256. <https://doi.org/10.1148/radiology.217.1.r00oc33251>
 76. Wormanns D, Kohl G, Klotz E, Marheine A, Beyer F, Heindel W, Diederich S (2004) Volumetric measurements of pulmonary nodules at multi-row detector CT: in vivo reproducibility. *Eur Radiol* 14(1):86–92. <https://doi.org/10.1007/s00330-003-2132-0>
 77. Yanagawa M, Tanaka Y, Leung AN, Morii E, Kusumoto M, Watanabe S, Watanabe H, Inoue M, Okumura M, Gyobu T, Ueda K, Honda O, Sumikawa H, Johkoh T, Tomiyama N (2014) Prognostic importance of volumetric measurements in stage I lung adenocarcinoma. *Radiology* 272(2):557–567. <https://doi.org/10.1148/radiol.14131903>
 78. Zhong Y, She Y, Deng J, Chen S, Wang T, Yang M, Ma M, Song Y, Qi H, Wang Y, Shi J, Wu C, Xie D, Chen C, Multi-omics Classifier for Pulmonary Nodules Collaborative G (2022) Deep learning for prediction of N2 metastasis and survival for clinical stage I non-small cell lung cancer. *Radiology* 302(1):200–211. <https://doi.org/10.1148/radiol.2021210902>
 79. Chen N, Zhou M, Dong X, Qu J, Gong F, Han Y, Qiu Y, Wang J, Liu Y, Wei Y, Xia J, Yu T, Zhang X, Zhang L (2020) Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study.

- Lancet 395(10223):507–513. [https://doi.org/10.1016/S0140-6736\(20\)30211-7](https://doi.org/10.1016/S0140-6736(20)30211-7)
80. Ito R, Iwano S, Naganawa S (2020) A review on the use of artificial intelligence for medical imaging of the lungs of patients with coronavirus disease 2019. *Diagn Interv Radiol* (Ankara, Turkey) 26(5):443–448
 81. Webb WR, Müller NL, Naidich DP (1992) High-resolution CT of the Lung, 4th edn. Wolters Kluwer, Lippincott Williams & Wilkins
 82. Kunihiro Y, Tanaka N, Kawano R, Yujiri T, Ueda K, Gondo T, Kobayashi T, Matsumoto T, Ito K (2022) High-resolution CT findings of pulmonary infections in patients with hematologic malignancy: comparison between patients with or without hematopoietic stem cell transplantation. *Jpn J Radiol* 40(8):791–799. <https://doi.org/10.1007/s11604-022-01260-7>
 83. Bernheim A, Mei X, Huang M, Yang Y, Fayad ZA, Zhang N, Diao K, Lin B, Zhu X, Li K, Li S, Shan H, Jacobi A, Chung M (2020) Chest CT findings in coronavirus disease-19 (COVID-19): relationship to duration of infection. *Radiology* 295(3):200463. <https://doi.org/10.1148/radiol.20200463>
 84. Wang Y, Dong C, Hu Y, Li C, Ren Q, Zhang X, Shi H, Zhou M (2020) Temporal changes of CT findings in 90 patients with COVID-19 pneumonia: a longitudinal study. *Radiology* 296(2):E55–E64. <https://doi.org/10.1148/radiol.20200843>
 85. Cozzi D, Cavigli E, Moroni C, Smorchkova O, Zantonelli G, Pradella S, Miele V (2021) Ground-glass opacity (GGO): a review of the differential diagnosis in the era of COVID-19. *Jpn J Radiol* 39(8):721–732. <https://doi.org/10.1007/s11604-021-01120-w>
 86. Aoki R, Iwasawa T, Hagiwara E, Komatsu S, Utsunomiya D, Ogura T (2021) Pulmonary vascular enlargement and lesion extent on computed tomography are correlated with COVID-19 disease severity. *Jpn J Radiol* 39(5):451–458. <https://doi.org/10.1007/s11604-020-01085-2>
 87. Zhu QQ, Gong T, Huang GQ, Niu ZF, Yue T, Xu FY, Chen C, Wang GB (2021) Pulmonary artery trunk enlargement on admission as a predictor of mortality in in-hospital patients with COVID-19. *Jpn J Radiol* 39(6):589–597. <https://doi.org/10.1007/s11604-021-01094-9>
 88. Fukuda A, Yanagawa N, Sekiya N, Ohyama K, Yomota M, Inui T, Fujiwara S, Kawai S, Fukushima K, Tanaka M, Kobayashi T, Yajima K, Imamura A (2021) An analysis of the radiological factors associated with respiratory failure in COVID-19 pneumonia and the CT features among different age categories. *Jpn J Radiol* 39(8):783–790. <https://doi.org/10.1007/s11604-021-01118-4>
 89. Kanayama A, Tsuchihashi Y, Otomi Y, Enomoto H, Arima Y, Takahashi T, Kobayashi Y, Kaku K, Sunagawa T, Suzuki M (2022) COVID-19 discharge summary database (CDS) group. Association of severe COVID-19 outcomes with radiological scoring and cardiomegaly: findings from the COVID-19 inpatients database, Japan. *Jpn J Radiol* 40(11):1138–1147. <https://doi.org/10.1007/s11604-022-01300-2>
 90. Inui S, Fujikawa A, Gonoi W, Kawano S, Sakurai K, Uchida Y, Ishida M, Abe O (2022) Comparison of CT findings of coronavirus disease 2019 (COVID-19) pneumonia caused by different major variants. *Jpn J Radiol* 40(12):1246–1256. <https://doi.org/10.1007/s11604-022-01301-1>
 91. Bai HX, Hsieh B, Xiong Z, Halsey K, Choi JW, Tran TML, Pan I, Shi LB, Wang DC, Mei J, Jiang XL, Zeng QH, Eglin TK, Hu PF, Agarwal S, Xie FF, Li S, Healey T, Atalay MK, Liao WH (2020) Performance of radiologists in differentiating COVID-19 from non-COVID-19 viral pneumonia at chest CT. *Radiology* 296(2):E46–E54. <https://doi.org/10.1148/radiol.20200823>
 92. Li L, Qin L, Xu Z, Yin Y, Wang X, Kong B, Bai J, Lu Y, Fang Z, Song Q, Cao K, Liu D, Wang G, Xu Q, Fang X, Zhang S, Xia J, Xia J (2020) Using artificial intelligence to detect COVID-19 and community-acquired pneumonia based on pulmonary CT: evaluation of the diagnostic accuracy. *Radiology* 296(2):E65–E71. <https://doi.org/10.1148/radiol.20200905>
 93. Roberts M, Driggs D, Thorpe M, Gilbey J, Yeung M, Ursprung S, Aviles-Rivero AI, Etmann C, McCague C, Beer L, Weir-McCall JR, Teng Z, Gkrania-Klotsas E, AIX-COVNET, Rudd JHF, Sala E, Schönlieb CB (2021) Common pitfalls and recommendations for using machine learning to detect and prognosticate for COVID-19 using chest radiographs and CT scans. *Nat Mach Intell* 3(3):199–217. <https://doi.org/10.1038/s42256-021-00307-0>
 94. Jia LL, Zhao JX, Pan NN, Shi LY, Zhao LP, Tian JH, Huang G (2022) Artificial intelligence model on chest imaging to diagnose COVID-19 and other pneumonias: a systematic review and meta-analysis. *Eur J Radiol Open* 9:100438. <https://doi.org/10.1016/j.ejro.2022.100438>
 95. Moroni C, Bindi A, Cavigli E, Cozzi D, Luvarà S, Smorchkova O, Zantonelli G, Miele V, Bartolucci M (2022) CT findings of non-neoplastic central airways diseases. *Jpn J Radiol* 40(2):107–119. <https://doi.org/10.1007/s11604-021-01190-w>
 96. Kunihiro Y, Tanaka N, Kawano R, Matsumoto T, Kobayashi T, Yujiri T, Kubo M, Gondo T, Ito K (2021) Differentiation of pulmonary complications with extensive ground-glass attenuation on high-resolution CT in immunocompromised patients. *Jpn J Radiol* 39(9):868–876. <https://doi.org/10.1007/s11604-021-01122-8>
 97. Goo HW, Park SH (2022) Prediction of pulmonary hypertension using central-to-peripheral pulmonary vascular volume ratio on three-dimensional cardiothoracic CT in patients with congenital heart disease. *Jpn J Radiol* 40(9):961–969. <https://doi.org/10.1007/s11604-022-01272-3>
 98. Goh LH, Tenant SC (2022) Incidental pulmonary emboli are associated with a very high probability of progressive malignant disease on staging CT scans. *Jpn J Radiol* 40(9):914–918. <https://doi.org/10.1007/s11604-022-01280-3>
 99. Javed QA, Sista AK (2019) Endovascular therapy for acute severe pulmonary embolism. *Int J Cardiovasc Imaging* 35(8):1443–1452. <https://doi.org/10.1007/s10554-019-01567-z>
 100. Agnelli G, Becattini C (2010) Acute pulmonary embolism. *N Engl J Med* 363(3):266–274. <https://doi.org/10.1056/NEJMr a0907731>
 101. Moores LK, Jackson WL Jr, Shorr AF, Jackson JL (2004) Meta-analysis: outcomes in patients with suspected pulmonary embolism managed with computed tomographic pulmonary angiography. *Ann Intern Med* 141(11):866–874. <https://doi.org/10.7326/0003-4819-141-11-200412070-00011>
 102. Anderson DR, Kahn SR, Rodger MA, Kovacs MJ, Morris T, Hirsch A, Lang E, Stiell I, Kovacs G, Dreyer J, Dennie C, Cartier Y, Barnes D, Burton E, Pleasance S, Skedgel C, O'Rourke K, Wells PS (2007) Computed tomographic pulmonary angiography vs ventilation-perfusion lung scanning in patients with suspected pulmonary embolism: a randomized controlled trial. *JAMA* 298(23):2743–2753. <https://doi.org/10.1001/jama.298.23.2743>
 103. Monti CB, Zanardo M, Cozzi A, Schiaffino S, Spagnolo P, Secchi F, De Cecco CN, Sardanelli F (2021) Dual-energy CT performance in acute pulmonary embolism: a meta-analysis. *Eur Radiol* 31(8):6248–6258. <https://doi.org/10.1007/s00330-020-07633-8>
 104. Soffer S, Klang E, Shimon O, Barash Y, Cahan N, Greenspan H, Konen E (2021) Deep learning for pulmonary embolism detection on computed tomography pulmonary angiogram: a systematic review and meta-analysis. *Sci Rep* 11(1):15814. <https://doi.org/10.1038/s41598-021-95249-3>
 105. Cheikh AB, Gorincour G, Nivet H, May J, Seux M, Calame P, Thomson V, Delabrousse E, Crombé A (2022) How artificial intelligence improves radiological interpretation in suspected pulmonary embolism. *Eur Radiol* 32(9):5831–5842. <https://doi.org/10.1007/s00330-022-08645-2>

106. Batra K, Xi Y, Al-Hreish KM, Kay FU, Browning T, Baker C, Peshock RM (2022) Detection of incidental pulmonary embolism on conventional contrast-enhanced chest CT: comparison of an artificial intelligence algorithm and clinical reports. *AJR Am J Roentgenol* 219(6):895–902. <https://doi.org/10.2214/AJR.22.27895>
107. Poggiali E, Bastoni D, Ioannilli E, Vercelli A, Magnacavallo A (2020) Deep vein thrombosis and pulmonary embolism: two complications of COVID-19 pneumonia? *Eur J Case Rep Intern Med* 7(5):001646. https://doi.org/10.12890/2020_001646
108. Meiler S, Hamer OW, Schaible J, Zeman F, Zorger N, Kleine H, Rennert J, Stroszczyński C, Poschenrieder F (2020) Computed tomography characterization and outcome evaluation of COVID-19 pneumonia complicated by venous thromboembolism. *PLoS ONE* 15(11):e0242475. <https://doi.org/10.1371/journal.pone.0242475>
109. Zaazoue KA, McCann MR, Ahmed AK, Cortopassi IO, Erben YM, Little BP, Stowell JT, Toskich BB, Ritchie CA (2023) Evaluating the performance of a commercially available artificial intelligence algorithm for automated detection of pulmonary embolism on contrast-enhanced computed tomography and computed tomography pulmonary angiography in patients with coronavirus disease 2019. *Mayo Clin Proc Innov Qual Outcomes* 3:143–152. <https://doi.org/10.1016/j.mayocpiqo.2023.03.001>
110. Dunnmon JA, Yi D, Langlotz CP, Ré C, Rubin DL, Lungren MP (2019) Assessment of convolutional neural networks for automated classification of chest radiographs. *Radiology* 290(2):537–544. <https://doi.org/10.1148/radiol.2018181422>
111. Maleki F, Ovens K, Gupta R, Reinhold C, Spatz A, Forghani R (2022) Generalizability of machine learning models: quantitative evaluation of three methodological pitfalls. *Radiol Artif Intell* 5(1):e220028. <https://doi.org/10.1148/ryai.220028>
112. Greffier J, Villani N, Defez D, Dabli D, Si-Mohamed S (2023) Spectral CT imaging: technical principles of dual-energy CT and multi-energy photon-counting CT. *Diagn Interv Imaging* 104(4):167–177. <https://doi.org/10.1016/j.diii.2022.11.003>
113. Nakamura Y, Higaki T, Kondo S, Kawashita I, Takahashi I, Awai K (2023) An introduction to photon-counting detector CT (PCD CT) for radiologists. *Jpn J Radiol* 41(3):266–282. <https://doi.org/10.1007/s11604-022-01350-6>
114. Bhayana R, Bleakney RR, Krishna S (2023) GPT-4 in radiology: improvements in advanced reasoning. *Radiology* 307(5):e230987. <https://doi.org/10.1148/radiol.230987>
115. Lourenco AP, Slanetz PJ, Baird GL (2023) Rise of ChatGPT: it may be time to reassess how we teach and test radiology residents. *Radiology* 307(5):e231053. <https://doi.org/10.1148/radiol.231053>
116. Khan WU, Seto E (2023) A “do no harm” novel safety checklist and research approach to determine whether to launch an artificial intelligence-based medical technology: introducing the biological-psychological, economic, and social (BPES) framework. *J Med Internet Res* 25:e43386. <https://doi.org/10.2196/43386>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.