QUALITY IN SPORT
eISSN 2450-3118 · Open Access · Peer-reviewed
apcz.umk.pl/QS Nicolaus Copernicus University in Toruń



CZERWONKA, Mikołaj, SOBIERAJSKI, Mateusz, SOCHAL, Wiktoria and KOT, Julia. Artificial intelligence in diagnostic imaging. Quality in Sport. 2026.79.76847. https://doi.org/10.12775/QS.2026.79.76847

ARTICLE TIMELINE
Received: 19.10.2026.
Accepted: 10.11.2026.
Published: 18.11.2026.

INDEXING & EVALUATION
MEiN points: 20 Unique ID: 201398
Disciplines: Economics & Finance; Management & Quality Sciences

The journal has been awarded 20 points in the parametric evaluation by the Ministry of Higher Education and Science of Poland. Annex to the announcement of the Minister of Higher Education and Science of 05.01.2024 No. 32553. Unique Journal Identifier: 201398. Scientific disciplines assigned: Economics and Finance (Field of Social Sciences); Management and Quality Sciences (Field of Social Sciences).
Punkty Ministerialne z 2019 — aktualny rok 20 punktów. Załącznik do komunikatu Ministra Szkolnictwa Wyższego i Nauki z dnia 05.01.2024 Lp. 32553. Posiada Unikatowy Identyfikator Czasopisma: 201398. Przypisane dyscypliny naukowe: Ekonomia i finanse (Dziedzina nauk społecznych); Nauki o zarządzaniu i jakości (Dziedzina nauk społecznych).

© The Authors 2026. This article is published with open access under the License Open Journal Systems of Nicolaus Copernicus University in Toruń, Poland. Open Access: This article is distributed under the terms of the Creative Commons Attribution Non-commercial Share Alike License (http://creativecommons.org/licenses/by-nc-sa/4.0/), which permits unrestricted, non-commercial use, distribution, and reproduction in any medium, provided the work is properly cited. The authors declare no conflict of interest regarding the publication of this paper.

Artificial Intelligence in Diagnostic Imaging

Wykorzystanie sztucznej inteligencji w diagnostyce obrazowej

Mikołaj Czerwonka

ORCID https://orcid.org/0009-0004-4494-1555

E- mail czerwonka.mikolaj88@gmail.com

Medical University of Lublin,

Aleje Racławickie 1, 20-059 Lublin, Poland

Mateusz Sobierajski

ORCID https://orcid.org/0009-0008-1859-9871

E- mail mateusz.sobierajski77@gmail.com

Medical University of Lublin,

Aleje Racławickie 1, 20-059 Lublin, Poland

Wiktoria Sochal
ORCID https://orcid.org/0009-0009-5234-1994
E- mail wsochal@gmail.com
Medical University of Lublin,
Aleje Racławickie 1, 20-059 Lublin, Poland

Julia Kot
ORCID https://orcid.org/0009-0005-6845-4177
E- mail kotjulia01@gmail.com
Medical University of Lublin,
Aleje Racławickie 1, 20-059 Lublin, Poland

Abstract

Background: Artificial intelligence is used to detect and characterize abnormalities in diagnostic imaging and to support clinical interpretation. Technical performance does not always translate into patient benefit.

Aim: To evaluate applications of artificial intelligence (AI) in diagnostic imaging, focusing on diagnostic performance, collaboration with clinicians, and limitations of clinical implementation.

Methods: To identify recent articles on the use of artificial intelligence in diagnostic imaging, the PubMed database was searched. Only English-language publications were included.

Results: AI algorithms have promising capabilities for detecting and characterizing abnormalities in mammography, radiography, computed tomography, magnetic resonance imaging, and ultrasonography. In selected settings, they improve sensitivity, support segmentation, and reduce interpretation workload. However, they do not consistently improve specificity, time to intervention, or patient outcomes. The clinical evidence varies in methodological quality and representativeness.

Conclusions: AI should primarily be considered a decision-support technology rather than an autonomous replacement for clinical judgment. Wider adoption requires external validation, workflow-based evaluation, and ongoing safety monitoring.
Keywords: artificial intelligence, diagnostic imaging, radiology, deep learning, computed tomography, magnetic resonance imaging, ultrasonography, mammography.

Streszczenie

Tło: Zastosowanie sztucznej inteligencji w diagnostyce obrazowej obejmuje wykrywanie i charakteryzowanie zmian oraz wspomaganie pracy lekarzy. Skuteczność techniczna algorytmów nie zawsze przekłada się na korzyści kliniczne.

Cel badań: Ocena zastosowań sztucznej inteligencji (AI) w diagnostyce obrazowej, ze szczególnym uwzględnieniem skuteczności diagnostycznej, współpracy z lekarzami i ograniczeń wdrożeń klinicznych.

Metody: W celu wyłonienia aktualnych artykułów dotyczących wykorzystania sztucznej inteligencji w diagnostyce obrazowej przeszukano bazę danych PubMed. Uwzględniono wyłącznie publikacje anglojęzyczne.

Wyniki: Algorytmy AI wykazują obiecujące możliwości wykrywania i charakteryzowania zmian w mammografii, RTG, TK, MRI i USG. W części zastosowań poprawiają czułość rozpoznania, wspomagają segmentację i ograniczają obciążenie odczytami. Nie zawsze jednak zwiększają swoistość, skracają czas do interwencji lub poprawiają wyniki leczenia. Dowody kliniczne różnią się jakością i reprezentatywnością.

Wnioski: AI należy rozpatrywać przede wszystkim jako narzędzie wspomagające decyzje medyczne. Jej upowszechnienie wymaga walidacji zewnętrznej, oceny w codziennym przepływie pracy i monitorowania bezpieczeństwa.
Słowa kluczowe: sztuczna inteligencja, diagnostyka obrazowa, radiologia, głębokie uczenie, tomografia komputerowa, rezonans magnetyczny, ultrasonografia, mammografia.

Introduction

Diagnostic imaging plays a fundamental role in diagnosing and monitoring many diseases. It includes conventional radiography (X-ray), computed tomography (CT), magnetic resonance imaging (MRI), ultrasonography (US), and specialized breast and cardiac imaging techniques. The growing number of examinations, the complexity of the resulting data, and staffing constraints have increased the need for tools that can make image analysis more efficient without compromising patient safety. In their review, Lawrence and colleagues described artificial intelligence as a technology that can support the detection of abnormalities, image interpretation, and workflow organization. At the same time, they emphasized that the evidence for benefits in real-world clinical settings remains inconclusive [1].

Artificial intelligence encompasses various computational approaches, including machine learning, deep learning, convolutional neural networks, and natural language processing. In image analysis, these methods are used, among other things, to detect, locate, segment, and classify abnormalities and to estimate disease risk. Tools supporting procedural planning, image reconstruction, and navigation have also been studied in interventional radiology. These applications extend beyond simply identifying abnormalities in an acquired image [2].

The performance of AI algorithms is assessed using measures such as sensitivity, specificity, and the area under the receiver operating characteristic curve (ROC). These measures describe classification performance, but they do not, by themselves, establish that introducing a system will reduce incorrect clinical decisions, unnecessary examinations, or delays in treatment [1] [3].

A Literature Review

Screening mammography is one of the best-studied applications of AI. The randomized MASAI trial, which involved more than 80,000 women, compared AI-supported image reading with standard assessment by two radiologists. In the AI group, 244 cancers were detected, compared with 203 in the control group, corresponding to 0.61% and 0.51% of examinations, respectively. However, the difference in cancer detection was insufficient to establish the superiority of either approach. False-positive results accounted for 1.5% in both groups, while the number of readings performed by radiologists decreased by 44.3%. Further examinations were recommended for 2.2% of women in the AI group and 2.0% in the control group. Among the detected cancers, 184 were invasive in the AI-supported group, compared with 165 after

standard double reading. However, this was a preliminary safety analysis, without complete follow-up of cancers diagnosed between screening rounds [5].

The possibility of replacing one of the two radiologist readings with an AI assessment was also examined in the ScreenTrustCAD study, involving 55,581 women. Combining the assessment of one radiologist with an independent AI reading detected 261 cases of breast cancer, whereas standard assessment by two radiologists identified 250. AI used alone identified 246 cancers, and a total of 269 cases were detected when the findings of all three readers were considered. These results confirmed that replacing one radiologist reading with AI did not reduce breast cancer detection in the screening program studied. All comparisons were based on the same mammograms, so the differences were not attributable to unequal numbers of patients in three separate groups [6].

However, not all studies have reported the same performance for standalone algorithms. A meta-analysis of 16 publications comprising more than 1.1 million mammographic and tomosynthesis assessments found that AI detected suspicious abnormalities more effectively than radiologists in some comparisons, but more often incorrectly classified images from healthy individuals as abnormal. Results also differed between studies in which clinicians and algorithms interpreted the same images and historical analyses of screening programs. The combined material included 497,091 women, although the algorithms were not all evaluated using the same imaging technique or under comparable conditions [7]. In practice, this means that detecting more abnormalities does not necessarily lead to fewer unnecessary additional examinations.

Different benefits were observed in a study of two groups of 1,579 mammograms each. Using AI to support interpretation increased the specificity of radiologists' assessments from 91.6% to 96.0%, while the proportion of examinations classified as abnormal decreased from 9.2% to 4.9%. At the same time, a similar number of cancers was detected in both groups, at 8.9 cases per 1,000 examinations [8]. Thus, in this application, the algorithm primarily helped reduce incorrect abnormal findings rather than increase the number of diagnosed cancers. A further review of 47 breast imaging studies, covering mammography, ultrasonography, and magnetic resonance imaging, confirmed that the effect of AI on sensitivity and specificity depends on the imaging method, the radiologist's experience, and how the system is used. In some publications, increased detection was accompanied by reduced specificity, while in others AI assistance produced no clear change in overall diagnostic accuracy [9].

Artificial intelligence has also been evaluated as a screening tool for tuberculosis on chest X-rays. In an analysis of certified products, the average sensitivity was 90.3% and specificity was 52.6%. This means that the algorithms identified a substantial proportion of people with tuberculosis but also relatively often flagged individuals in whom the disease was not confirmed. Differences between products were considerable: qXR achieved a sensitivity of 94.4% and a specificity of 69.2%, Lunit INSIGHT achieved 85.3% and 64.6%, respectively, and CAD4TB achieved 91.7% and 37.1%. A positive image-analysis result therefore could not replace further microbiological testing. The review included 10 studies, with data from eight used in the pooled analysis. Such large differences in specificity show that the choice of product affects how many people are referred for confirmatory testing [10].

In chest computed tomography, an important task is not only to detect a pulmonary nodule but also to determine whether it may be malignant. A review of 17 externally validated studies analyzed data from 8,553 individuals and 9,884 pulmonary nodules. Deep learning models showed greater sensitivity in identifying malignant lesions than unaided clinical assessment, with comparable specificity of 77% for AI and 81% for clinicians. The comparison concerned the risk of malignancy in nodules already detected on CT, rather than the automatic identification of all lung abnormalities [11]. Comparisons of CT readings performed with and without algorithmic assistance also showed that AI can increase nodule detection, although sometimes at the cost of more false-positive results. Assistance yielded more favorable results when assessing the malignancy of previously detected lesions. It was therefore important to distinguish between using an algorithm to locate a small lesion and using it to help a clinician assess a nodule that had already been identified [12].

Detecting more abnormalities does not always translate into a better subsequent diagnostic process. An assessment of 27 studies of software for analyzing pulmonary nodules on CT found that these systems could increase detection and shorten image interpretation time, but might also lead to more frequent referrals for follow-up CT examinations. All included studies were judged to have a high risk of bias, and the available data did not clearly demonstrate improvements in long-term health outcomes or the cost-effectiveness of implementation [13].

Another application is the detection of pulmonary embolism on CT pulmonary angiography. A review of 12 studies covered more than 341,000 analyzed images in total. One of the evaluated models detected approximately 88–91% of embolism cases, although small abnormalities in the pulmonary vessels were particularly difficult to identify. The limited number of studies conducted outside the centers where the algorithms were developed made it difficult to

determine their performance in routine clinical practice. In some studies, false-positive findings occurred even when embolism detection was good, potentially increasing the number of images requiring additional review by a physician [14].

In neuroradiology, AI is used, among other applications, to detect brain tumors on magnetic resonance images. An analysis of 19 studies involving approximately 12,000 patients found a sensitivity of 87% for tumor detection at the patient level and 86% at the level of individual lesions. Algorithms also helped automatically delineate tumor boundaries, although the quality of these results varied according to tumor type, MRI scanner, and image-processing method. In some studies, models analyzing three-dimensional images achieved better detection results than two-dimensional approaches. However, the accuracy of tumor boundary delineation was not equivalent to the ability to detect the tumor itself [15].

The use of AI for detecting intracranial hemorrhage has also been studied in brain CT. A meta-analysis of 36 publications found an average sensitivity of 89% and specificity of 91%. Good results were also obtained when automatically delineating the area of hemorrhage, which may facilitate assessment of its extent. The positive predictive value was 92% and the negative predictive value was 94%, indicating favorable results both for identifying suspected hemorrhages and for ruling them out in the samples assessed [16]. These figures demonstrate good performance in identifying abnormalities on images, but do not in themselves establish that treatment can begin sooner.

The importance of this distinction is evident in studies of systems supporting acute stroke management. Software analyzing brain CT can assist in identifying ischemic changes, vessel occlusions, and cases requiring urgent assessment. However, a review of the available studies did not provide sufficient evidence that using such tools improves treatment outcomes or produces definite cost savings [17]. Similarly, in an analysis of 2,306 patients, an AI system for intracranial hemorrhage detection flagged 288 examinations as positive, and 82.3% of these findings were correct. Despite automatically prioritizing suspicious images, the system did not produce a significant reduction in time to urgent intervention [18].

Good diagnostic performance has also been reported for the automatic detection of fractures on X-rays. A meta-analysis of 100 studies found that the average sensitivity of AI algorithms was approximately 91% and specificity approximately 92%. AI performance was similar to that of clinicians, although experienced radiologists were less likely to incorrectly diagnose a fracture in some comparisons. Results also varied by injury location, so a single performance estimate

cannot be applied to every bone. Detailed data suitable for estimating performance were available from 83 of the 100 included publications. Separate assessments of individual anatomical sites indicated that the difficulty of detecting a fracture also depends on its appearance and location [19].

Similar findings were reported in a review of 66 studies on fracture detection. Image-based models achieved a sensitivity of 92% and specificity of 91%, whereas systems using tabular data alone achieved 81% and 83%, respectively. For hip fracture detection, sensitivity was 92% and specificity was 90%. Spinal and hip imaging were the most frequently studied areas, with 20 and 18 studies, respectively, although these two locations do not represent the full range of injuries [20]. Together, these findings indicate that direct image analysis is particularly important in trauma diagnosis, while performance still depends on the type of fracture being assessed.

Not all commercially available systems achieve similar results. A review of 17 studies evaluating seven commercial products included 38,978 X-rays, of which 8,150 showed fractures. Four of the five systems assessed independently detected more than 90% of fractures, whereas one detected fewer than 60%. Rib and spinal fractures proved particularly difficult to assess. In some studies, combining clinical interpretation with AI improved results compared with the algorithm operating alone. Some studies of products funded by their manufacturers reported better detection performance, highlighting the importance of independent evaluation of commercial systems [21].

The diagnosis of hip fractures and the prediction of outcomes after surgery were also evaluated separately. In a review of 39 studies, 18 examined fracture detection and 21 addressed the prediction of postoperative outcomes. The average sensitivity of the diagnostic models was 89.3% and specificity was 87.5%. However, no clear advantage of AI over clinicians was demonstrated in terms of incorrect diagnoses. Similarly, for predicting outcomes after surgery, more advanced algorithms did not achieve markedly better results than traditional statistical methods [22].

In ultrasonography, artificial intelligence is used primarily to distinguish benign from malignant lesions. A review of 52 studies involving focal liver lesions found that some models achieved performance comparable to that of experienced physicians. However, the studies differed in lesion type, image quality, and methods used to verify their results, making it difficult to identify a single level of performance across all applications. This applied both to conventional ultrasound imaging and to contrast-enhanced ultrasound [23].

In contrast-enhanced liver ultrasonography, 41 studies were analyzed, with reported lesion classification accuracy ranging widely from 58.6% to 98.9%. In 25 studies, the algorithms analyzed images that had already been selected or annotated by a person. High performance on individual frames may therefore not translate into equally good results when interpreting an entire examination in real time [24]. These differences demonstrate the importance of how imaging data are prepared.

The potential of AI to support cancer diagnosis has also been assessed in ovarian ultrasonography. A meta-analysis of 14 studies covering 15,358 images found a sensitivity of 81% and specificity of 92% for distinguishing malignant from nonmalignant lesions, using histopathological examination as the reference standard [25]. In pancreatic imaging, researchers have examined algorithms capable of identifying subtle changes on CT even before a clinical cancer diagnosis. The available studies suggested that these systems could detect suspicious features earlier and assist with segmentation and disease staging. However, they did not establish a single reliable performance estimate for this entire group of applications or confirm their performance in prospective studies [26].

In cardiac imaging, AI is used for image reconstruction, identification of anatomical structures, automated measurements, and diagnostic support. A statement from international scientific societies emphasized that individual solutions are at different stages of implementation, ranging from experimental tools to those already used clinically. Their usefulness should therefore be evaluated separately for computed tomography, magnetic resonance imaging, and specific diagnostic tasks. Particular attention is needed regarding the safety of image reconstruction, measurement quality, and the protection of patient data [27].

One of the larger studies of cardiac magnetic resonance imaging developed a system for screening and diagnosing 11 groups of cardiovascular diseases. The screening stage was evaluated using data from 9,719 patients and the diagnostic stage using data from 8,066 individuals. The algorithm achieved very high performance in distinguishing the diseases studied and outperformed the cardiologists included in the comparison when identifying pulmonary arterial hypertension. The data came from eight centers using scanners from three manufacturers, although these results cannot be generalized to all AI systems used in cardiology [28].

In echocardiography, algorithms assist with recognizing imaging views, delineating cardiac chambers, and performing measurements that would normally require manual analysis by a

clinician. A review of available studies indicated that these tools may reduce differences between observers and partially automate interpretation. However, the results depended on the quality of the recorded images, and no single performance measure could be provided for all echocardiographic applications. Unlike detecting a single abnormality, measuring cardiac structures requires reproducible results across different views and examination conditions [29].

A comparison of the publications discussed shows that the benefits of artificial intelligence vary according to the purpose of the examination. In mammography, detecting cancer while limiting unnecessary recalls is important; in fracture imaging, the accurate diagnosis of an injury is the main goal; and in brain CT, the speed of reporting findings and initiating treatment also matters. For this reason, good image-based performance should not be equated with a demonstrated improvement in patient outcomes [5][18][19].

False-positive results also remain an important limitation. In tuberculosis screening, high sensitivity was accompanied by low specificity in some systems, while in pulmonary nodule assessment, increased detection could lead to more frequent follow-up of benign lesions [10][13]. In addition, incomplete reporting of model development methods, insufficient external validation, and differences between equipment manufacturers make it difficult to predict an algorithm's performance in a new center [4][15][24].

Current findings support the use of AI as an aid to clinicians in selected diagnostic tasks. Before wider implementation, however, further studies are needed to assess not only the accuracy of abnormality detection but also the number of unnecessary examinations, time to treatment, patient safety, and the costs of the entire diagnostic pathway [1][3][17][27].

Conclusions

Artificial intelligence is highly useful in selected diagnostic imaging tasks, particularly the detection and classification of abnormalities in mammography, X-ray, CT, MRI, and ultrasonography. However, performance depends on the examination type, patient population, reference standard, and specific algorithm [7][15][16][19][24][25][28].

In selected settings, AI assistance can increase the sensitivity of clinical assessment, improve the reproducibility of measurements, and reduce the number of readings performed by radiologists. Prospective studies of screening mammography provide particularly valuable evidence of feasibility, but their results should not automatically be generalized to other fields [5][6][8][12][21].

A high proportion of correctly identified abnormalities does not automatically translate into improved treatment outcomes. In one study, implementing a system for detecting intracranial

hemorrhage did not significantly shorten the time to intervention, while in pulmonary imaging, increased detection may be accompanied by more follow-up examinations [13][18].

The main limitations of the current evidence include methodological heterogeneity, insufficient external validation, incomplete reporting of models, the risk of false-positive findings, and a lack of adequate data on long-term clinical consequences and costs [1][4][13][23][24].

AI should primarily be regarded as a tool that supports clinicians. Its safe and widespread adoption requires multicenter prospective studies, assessment of its impact on the entire diagnostic process, monitoring of errors, and a clear definition of responsibility for medical decisions [2][3][9][27].

Author Contributions

Conceptualization, Mikołaj Czerwonka ; methodology, Mikołaj Czerwonka; investigation, Julia Kot, and Wiktoria Sochal; resources, Mateusz Sobierajski; validation, Mateusz Sobierajski; visualization, Wiktoria Sochal; writing- original draft preparation, Mikołaj Czerwonka, Julia Kot, Wiktoria Sochal and Mateusz Sobierajski; writing – review and editing, Mikołaj Czerwonka and Mateusz Sobierajski; project administration, Mikołaj Czerwonka

All authors have read and agreed to the published version of the manuscript.

Funding

This research received no external funding.

Institutional Review Board Statement

Not applicable.

Informed Consent Statement

Not applicable.

Data Availability Statement

Not applicable.

Conflict of interest

The authors declare no conflict of interest.

Declaration of AI and AI- assisted Technologies in the Writing Process

The author used ChatGPT solely for the purpose of translating a small portion of the manuscript into English. The translated text was subsequently reviewed and approved by the authors, who take full responsibility for the scientific content of the publication.

References

[1] Lawrence R, Dodsworth E, Massou E, Sherlaw-Johnson C, Ramsay AIG, Walton H, O'Regan T, Gleeson F, Crellin N, Herbert K, Ng PL, Elphinstone H, Mehta R, Lloyd J, Halliday A, Morris S, Fulop NJ. Artificial intelligence for diagnostics in radiology practice: a rapid systematic scoping review. EClinicalMedicine. 2025 May 12;83:103228. doi: 10.1016/j.eclinm.2025.103228. PMID: 40474995; PMCID: PMC12140059.

[2] Singh SP, Ramprasad A, Makary MS. Clinical utility of artificial intelligence models in radiology: a systemic scoping review of diagnostic and endovascular applications. CVIR Endovasc. 2025 Oct 30;8(1):97. doi: 10.1186/s42155-025-00573-8. PMID: 41162783; PMCID: PMC12572426.

[3] Wenderott K, Krups J, Zaruchas F, Weigl M. Effects of artificial intelligence implementation on efficiency in medical imaging—a systematic literature review and meta-analysis. NPJ Digit Med. 2024 Sep 30;7(1):265. doi: 10.1038/s41746-024-01248-9. PMID: 39349815; PMCID: PMC11442995.

[4] Moassefi M, Rouzrokh P, Conte GM, Vahdati S, Fu T, Tahmasebi A, Younis M, Farahani K, Gentili A, Kline T, Kitamura FC, Huo Y, Kuanar S, Younis K, Erickson BJ, Faghani S. Reproducibility of Deep Learning Algorithms Developed for Medical Imaging Analysis: A Systematic Review. J Digit Imaging. 2023 Oct;36(5):2306-2312. doi: 10.1007/s10278-023-00870-5. PMID: 37407841; PMCID: PMC10501962.

[5] Lång K, Josefsson V, Larsson AM, Larsson S, Högberg C, Sartor H, Hofvind S, Andersson I, Rosso A. Artificial intelligence-supported screen reading versus standard double reading in the Mammography Screening with Artificial Intelligence trial (MASAI): a clinical safety analysis of a randomised, controlled, non-inferiority, single-blinded, screening accuracy study. Lancet Oncol. 2023 Aug;24(8):936-944. doi: 10.1016/S1470-2045(23)00298-X. PMID: 37541274.

[6] Dembrower K, Crippa A, Colón E, Eklund M, Strand F; ScreenTrustCAD Trial Consortium. Artificial intelligence for breast cancer detection in screening mammography in Sweden: a

prospective, population-based, paired-reader, non-inferiority study. Lancet Digit Health. 2023 Oct;5(10):e703-e711. doi: 10.1016/S2589-7500(23)00153-X. PMID: 37690911.

[7] Yoon JH, Strand F, Baltzer PAT, Conant EF, Gilbert FJ, Lehman CD, Morris EA, Mullen LA, Nishikawa RM, Sharma N, Vejborg I, Moy L, Mann RM. Standalone AI for Breast Cancer Detection at Screening Digital Mammography and Digital Breast Tomosynthesis: A Systematic Review and Meta-Analysis. Radiology. 2023 Jun;307(5):e222639. doi: 10.1148/radiol.222639. PMID: 37219445; PMCID: PMC10315526.

[8] Kim H, Choi JS, Kim K, Ko ES, Ko EY, Han BK. Effect of artificial intelligence-based computer-aided diagnosis on the screening outcomes of digital mammography: a matched cohort study. Eur Radiol. 2023 Oct;33(10):7186-7198. doi: 10.1007/s00330-023-09692-z. PMID: 37188881.

[9] Lu C, Tan T, Yu L, MohammadiNasab P, Zhang T, Zhang H, Soryani M, Mann R, Kozegar E, Bao L. AI in breast imaging: A systematic review of reader studies evaluating design, performance, and clinical impact. Med Phys. 2025 Oct;52(10):e70018. doi: 10.1002/mp.70018. PMID: 41065282.

[10] Hua D, Nguyen K, Petrina N, Young N, Cho JG, Yap A, Poon SK. Benchmarking the diagnostic test accuracy of certified AI products for screening pulmonary tuberculosis in digital chest radiographs: Preliminary evidence from a rapid review and meta-analysis. Int J Med Inform. 2023 Sep;177:105159. doi: 10.1016/j.ijmedinf.2023.105159. PMID: 37549498.

[11] Wulaningsih W, Villamaria C, Akram A, Benemile J, Croce F, Watkins J. Deep Learning Models for Predicting Malignancy Risk in CT-Detected Pulmonary Nodules: A Systematic Review and Meta-analysis. Lung. 2024 Oct;202(5):625-636. doi: 10.1007/s00408-024-00706-1. PMID: 38782779; PMCID: PMC11427562.

[12] Ewals LJS, van der Wulp K, van den Borne BEEM, Pluyter JR, Jacobs I, Mavroeidis D, van der Sommen F, Nederend J. The Effects of Artificial Intelligence Assistance on the Radiologists' Assessment of Lung Nodules on CT Scans: A Systematic Review. J Clin Med.

2023 May 18;12(10):3536. doi: 10.3390/jcm12103536. PMID: 37240643; PMCID: PMC10219568.

[13] Geppert J, Auguste P, Asgharzadeh A, Ghiasvand H, Patel M, Brown A, Jayakody S, Helm E, Todkill D, Madan J, Stinton C, Gallacher D, Taylor-Phillips S, Chen YF. Software with artificial intelligence-derived algorithms for detecting and analysing lung nodules in CT scans: systematic review and economic evaluation. Health Technol Assess. 2025 May;29(14):1-234. doi: 10.3310/JYTW8921. PMID: 40380885; PMCID: PMC12104856.

[14] Hassan MSTA, Elhotiby MAM, Shah V, Rocha H, Arjomandi Rad A, Miller G, Malawana J. The Current State of Artificial Intelligence on Detecting Pulmonary Embolism via Computerised Tomography Pulmonary Angiogram: A Systematic Review. Br J Hosp Med (Lond). 2025 Jun 25;86(6):1-21. doi: 10.12968/hmed.2024.0757. PMID: 40554436.

[15] Wang TW, Shiao YC, Hong JS, Lee WK, Hsu MS, Cheng HM, Yang HC, Lee CC, Pan HC, You WC, Lirng JF, Guo WY, Wu YT. Artificial Intelligence Detection and Segmentation Models: A Systematic Review and Meta-Analysis of Brain Tumors in Magnetic Resonance Imaging. Mayo Clin Proc Digit Health. 2024 Feb 4;2(1):75-91. doi: 10.1016/j.mcpdig.2024.01.002. PMID: 40206681; PMCID: PMC11976016.

[16] Hu P, Yan T, Xiao B, Shu H, Sheng Y, Wu Y, Shu L, Lv S, Ye M, Gong Y, Wu M, Zhu X. Deep learning-assisted detection and segmentation of intracranial hemorrhage in noncontrast computed tomography scans of acute stroke patients: a systematic review and meta-analysis. Int J Surg. 2024 Jun 1;110(6):3839-3847. doi: 10.1097/JS9.0000000000001266. PMID: 38489547; PMCID: PMC11175741.

[17] Westwood M, Ramaekers B, Grimm S, Armstrong N, Wijnen B, Ahmadu C, de Kock S, Noake C, Joore M. Software with artificial intelligence-derived algorithms for analysing CT brain scans in people with a suspected acute stroke: a systematic review and cost-effectiveness analysis. Health Technol Assess. 2024 Mar;28(11):1-204. doi: 10.3310/RDPA1487. PMID: 38512017; PMCID: PMC11017149.

[18] Bark D, Basu J, Toumpanakis D, Nyberg JB, Bjerner T, Rostami E, Fällmar D. Clinical Impact of an AI Decision Support System for Detection of Intracranial Hemorrhage in CT

Scans. Neurotrauma Rep. 2024 Oct 14;5(1):1009-1015. doi: 10.1089/neur.2024.0017. PMID: 39440151; PMCID: PMC11491571.

[19] Nowroozi A, Salehi MA, Shobeiri P, Agahi S, Momtazmanesh S, Kaviani P, Kalra MK. Artificial intelligence diagnostic accuracy in fracture detection from plain radiographs and comparing it with clinicians: a systematic review and meta-analysis. Clin Radiol. 2024 Aug;79(8):579-588. doi: 10.1016/j.crad.2024.04.009. PMID: 38772766.

[20] Jung J, Dai J, Liu B, Wu Q. Artificial intelligence in fracture detection with different image modalities and data types: A systematic review and meta-analysis. PLOS Digit Health. 2024 Jan 30;3(1):e0000438. doi: 10.1371/journal.pdig.0000438. PMID: 38289965; PMCID: PMC10826962.

[21] Husarek J, Hess S, Razaeian S, Ruder TD, Sehmisch S, Müller M, Liodakis E. Artificial intelligence in commercial fracture detection products: a systematic review and meta-analysis of diagnostic test accuracy. Sci Rep. 2024 Oct 4;14(1):23053. doi: 10.1038/s41598-024-73058-8. PMID: 39367147; PMCID: PMC11452402.

[22] Lex JR, Di Michele J, Koucheki R, Pincus D, Whyne C, Ravi B. Artificial Intelligence for Hip Fracture Detection and Outcome Prediction: A Systematic Review and Meta-analysis. JAMA Netw Open. 2023 Mar 1;6(3):e233391. doi: 10.1001/jamanetworkopen.2023.3391. PMID: 36930153; PMCID: PMC10024206.

[23] Vetter M, Waldner MJ, Zundler S, Klett D, Bocklitz T, Neurath MF, Adler W, Jesper D. Artificial intelligence for the classification of focal liver lesions in ultrasound – a systematic review. Ultraschall Med. 2023 Aug;44(4):395-407. doi: 10.1055/a-2066-9372. PMID: 37001563.

[24] Brooks JA, Kallenbach M, Radu IP, Berzigotti A, Dietrich CF, Kather JN, Luedde T, Seraphin TP. Artificial Intelligence for Contrast-Enhanced Ultrasound of the Liver: A Systematic Review. Digestion. 2025;106(3):227-244. doi: 10.1159/000541540. PMID: 39312896; PMCID: PMC12129427.

[25] Mitchell S, Nikolopoulos M, El-Zarka A, Al-Karawi D, Al-Zaidi S, Ghai A, Gaughran JE, Sayasneh A. Artificial Intelligence in Ultrasound Diagnoses of Ovarian Cancer: A Systematic Review and Meta-Analysis. Cancers (Basel). 2024 Jan 19;16(2):422. doi: 10.3390/cancers16020422. PMID: 38275863; PMCID: PMC10813993.

[26] Antony A, Mukherjee S, Bhinder K, Murlidhar M, Zarrintan A, Goenka AH. Artificial Intelligence-Augmented Imaging for Early Pancreatic Cancer Detection. Visc Med. 2025 Oct;41(5):271-279. doi: 10.1159/000546603. PMID: 40567392; PMCID: PMC12187166.

[27] Mastrodicasa D, van Assen M, Huisman M, Leiner T, Williamson EE, Nicol ED, Allen BD, Saba L, Vliegenthart R, Hanneman K. Use of AI in Cardiac CT and MRI: A Scientific Statement from the ESCR, EuSoMII, NASCI, SCCT, SCMR, SIIM, and RSNA. Radiology. 2025 Jan;314(1):e240516. doi: 10.1148/radiol.240516. PMID: 39873607; PMCID: PMC11783164.

[28] Wang YRJ, Yang K, Wen Y, Wang P, Hu Y, Lai Y, Wang Y, Zhao K, Tang S, Zhang A, Zhan H, Lu M, Chen X, Yang S, Dong Z, Wang Y, Liu H, Zhao L, Huang L, Li Y, Wu L, Chen Z, Luo Y, Liu D, Zhao P, Lin K, Wu JC, Zhao S. Screening and diagnosis of cardiovascular disease using artificial intelligence-enabled cardiac magnetic resonance imaging. Nat Med. 2024 May;30(5):1471-1480. doi: 10.1038/s41591-024-02971-2. PMID: 38740996; PMCID: PMC11108784.

[29] Barry T, Farina JM, Chao CJ, Ayoub C, Jeong J, Patel BN, Banerjee I, Arsanjani R. The Role of Artificial Intelligence in Echocardiography. J Imaging. 2023 Feb 20;9(2):50. doi: 10.3390/jimaging9020050. PMID: 36826969; PMCID: PMC9962859.