Review

Artificial Intelligence Performance in Image-Based Cancer Identification: Umbrella Review of Systematic Reviews

He-Li Xu[1,2,3*], MPH; Ting-Ting Gong[4*], Prof Dr Med; Xin-Jian Song[1,2,3*], MPH; Qian Chen[4], MBBS; Qi Bao[4,5], MPH; Wei Yao[1,2,3], MPH; Meng-Meng Xie[4], MM; Chen Li[6], Prof Dr; Marcin Grzegorzek[7], Prof Dr; Yu Shi[8], Prof Dr Med; Hong-Zan Sun[8], Prof Dr Med; Xiao-Han Li[9], Prof Dr Med; Yu-Hong Zhao[1,2,3], Prof Dr Med; Song Gao[4], Prof Dr Med; Qi-Jun Wu[2,3,4,10,11], Prof Dr Med

[1]Department of Clinical Epidemiology, Shengjing Hospital of China Medical University, Shenyang, China

[2]Clinical Research Center, Shengjing Hospital of China Medical University, Shenyang, China

[3]Liaoning Key Laboratory of Precision Medical Research on Major Chronic Disease, Shengjing Hospital of China Medical University, Shenyang, China

[4]Department of Obstetrics and Gynecology, Shengjing Hospital of China Medical University, Shenyang, China

[5]Department of Epidemiology, School of Public Health, China Medical University, Shenyang, China

[6]Microscopic Image and Medical Image Analysis Group, College of Medicine and Biological Information Engineering, Northeastern University, Shenyang, China

[7]Institute for Medical Informatics, University of Luebeck, Luebeck, Germany

[8]Department of Radiology, Shengjing Hospital of China Medical University, Shenyang, China

[9]Department of Pathology, Shengjing Hospital of China Medical University, Shenyang, China

[10]Department of Clinical Epidemiology, Shengjing Hospital of China Medical University, Shenyang, Liaoning, China

[11]NHC Key Laboratory of Advanced Reproductive Medicine and Fertility (China Medical University), National Health Commission, Shenyang, China

*these authors contributed equally

Corresponding Author:
Qi-Jun Wu, Prof Dr Med
Department of Clinical Epidemiology
Shengjing Hospital of China Medical University
No. 36, San Hao Street
Shenyang, Liaoning, 110004
China
Phone: 86 024 96615 13652
Email: wuqj@sj-hospital.org

Abstract

Background: Artificial intelligence (AI) has the potential to transform cancer diagnosis, ultimately leading to better patient outcomes.

Objective: We performed an umbrella review to summarize and critically evaluate the evidence for the AI-based imaging diagnosis of cancers.

Methods: PubMed, Embase, Web of Science, Cochrane, and IEEE databases were searched for relevant systematic reviews from inception to June 19, 2024. Two independent investigators abstracted data and assessed the quality of evidence, using the Joanna Briggs Institute (JBI) Critical Appraisal Checklist for Systematic Reviews and Research Syntheses. We further assessed the quality of evidence in each meta-analysis by applying the Grading of Recommendations, Assessment, Development, and Evaluation (GRADE) criteria. Diagnostic performance data were synthesized narratively.

Results: In a comprehensive analysis of 158 included studies evaluating the performance of AI algorithms in noninvasive imaging diagnosis across 8 major human system cancers, the accuracy of the classifiers for central nervous system cancers varied widely (ranging from 48% to 100%). Similarities were observed in the diagnostic performance for cancers of the head and neck, respiratory system, digestive system, urinary system, female-related systems, skin, and other sites. Most meta-analyses demonstrated positive summary performance. For instance, 9 reviews meta-analyzed sensitivity and specificity for esophageal cancer, showing ranges of 90%-95% and 80%-93.8%, respectively. In the case of breast cancer detection, 8 reviews calculated the pooled sensitivity and specificity within the ranges of 75.4%-92% and 83%-90.6%, respectively. Four meta-analyses reported the ranges of sensitivity

and specificity in ovarian cancer, and both were 75%-94%. Notably, in lung cancer, the pooled specificity was relatively low, primarily distributed between 65% and 80%. Furthermore, 80.4% (127/158) of the included studies were of high quality according to the JBI Critical Appraisal Checklist, with the remaining studies classified as medium quality. The GRADE assessment indicated that the overall quality of the evidence was moderate to low.

Conclusions: Although AI shows great potential for achieving accelerated, accurate, and more objective diagnoses of multiple cancers, there are still hurdles to overcome before its implementation in clinical settings. The present findings highlight that a concerted effort from the research community, clinicians, and policymakers is required to overcome existing hurdles and translate this potential into improved patient outcomes and health care delivery.

Trial Registration: PROSPERO CRD42022364278; https://www.crd.york.ac.uk/PROSPERO/view/CRD42022364278

Introduction

Globally, cancer accounted for 9.7 million deaths in 2022, and it accounts for nearly 1 in 6 deaths, based on updated estimates from the International Agency for Research on Cancer [1]. It is estimated that the global annual number of new cancer cases will increase to 35 million by 2050, causing a major global burden that varies markedly across countries and territories [1,2]. Due to its disordered growth, unlimited proliferation, and easy metastasis, patients at the terminal stage of cancer show a high probability of death [3]. The sensitive and accurate detection of cancers is critical to realizing timely and effective treatment, which can improve the survival rate of patients and reduce their pain [4].

In modern clinical practice, biomedical imaging systems represent one of the main pillars of comprehensive cancer diagnosis and grading [5]. Traditionally, pathological analysis has been considered essential to determine the malignancy of a tumor as it involves image analysis at the cellular level [6]. Additionally, noninvasive imaging techniques, such as ultrasound (US), computed tomography (CT), magnetic resonance imaging (MRI), and positron emission tomography (PET), have been popularly employed for tumor diagnosis [7]. However, biomedical imaging assessment most commonly relies upon visual evaluation, which is a subjective, susceptible, time-consuming, and labor-intensive process [8]. The interpretation of imaging findings may be augmented by advanced computational analyses [9]. With the high rate of production of images and the increasing reliance on images, the use of artificial intelligence (AI) to assist in imaging diagnosis has been a growing trend and an active research field [10,11]. Computer-aided diagnosis systems for medical imaging involving AI could help to overcome the mismatch between the increasing number of images and the capacity of available specialists [12].

Over the past decade, AI has made substantial strides in medical imaging, allowing machines to better interpret complex data [9]. In simple terms, AI excels at recognizing complex patterns in images and thus offers the opportunity to transform image assessment from a purely qualitative and subjective task to one that is quantifiable and effortlessly reproducible [9,13]. This property of AI extends its applicability to accurately detecting cancer from medical images, thereby holding the potential to unburden clinicians from repetitive tasks [14]. In the past, numerous studies have been published on AI-based imaging analyses in the diagnosis of different cancers. Advances in the use of AI for cancer diagnostic research are on the rise, with multiple examples, including the successful classification of dermoscopy images [15], the interpretation of mammograms for breast cancer screening [16], and the image-based differentiation between ovarian cancer and its benign imitators [17]. However, previous efforts to systematically appraise the evidence have been focused on a single tumor.

To assimilate the vast amount of research available and accelerate clinical transformation, we conducted an overview of systematic reviews for the purpose of systematically synthesizing previously published evidence on the performance of AI algorithms in the noninvasive imaging diagnosis of cancer.

Methods

Guidelines and Registration

This umbrella review was performed in line with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines [18] and Joanna Briggs Institute (JBI) guidelines (Multimedia Appendix 1) [19]. The predefined protocol was registered in the international open-access Prospective Register of Systematic Reviews (PROSPERO; CRD42022364278).

Data Sources and Search Strategy

The present systematic literature search was completed using the following databases: PubMed, Embase, Web of Science, Cochrane, and IEEE. The search terms were as follows: ("artificial intelligence" OR "machine learning" OR "deep learning" OR "neural network") AND ("carcinoma" OR "tumor" OR "cancer" OR "neoplas*" OR "maligna*") AND ("meta-analysis" OR "systematic review"). The search was limited to articles written in English and published from inception to June 19, 2024. For the full search strategy, please see Multimedia Appendix 2.

Eligibility Criteria

All systematic reviews with or without a meta-analysis that examined the noninvasive imaging diagnostic performance of

AI technologies for cancers were eligible for inclusion. We also included reviews that had a systematic strategy for the literature search, even if the authors did not specify the article type [20]. Studies were excluded if they (1) focused on AI-based approaches for predicting the outcomes of interventions or the prognosis of cancers; (2) did not show at least one of the following measures of classifier performance: accuracy, sensitivity, specificity, or area under the curve (AUC); (3) did not involve noninvasive images, a systematic review, or interested outcomes; (4) were primary studies, scoping reviews, literature reviews, rapid reviews, criterial reviews, and other types of reviews; and (5) were conference abstracts and posters, commentaries, preprints, proposals, and editorials.

Selection Process

Following the removal of duplicates, 2 investigators (TTG and HLX) independently screened the titles and abstracts to identify potentially relevant articles. The full texts of these articles were then assessed for eligibility based on the inclusion and exclusion criteria. Disagreement among the 2 reviewers was resolved by discussion with a third reviewer (XJS).

Data Extraction

Two investigators (XJS and QB) independently extracted valuable data, using a predefined data extraction sheet. The following data were extracted: (1) characteristics of eligible reviews, including first author, year of publication, country, cancer site, registered protocol (yes or no), followed guidelines, research question, database searched, language restrictions, numbers of retrieved studies and included studies, number of reviewers (study selection, data extraction, and quality assessment), quality assessment tool, and meta-analysis (yes or no); (2) features of models in the included reviews, including imaging modalities, dataset size, AI approach, classification algorithms, and type of validation; and (3) classifier performance in multiple system cancers, including accuracy (n), sensitivity (n), specificity (n), AUC (n), and I^2. Points of divergence were resolved by discussion among 2 authors (HLX and TTG).

Quality Assessment and the Certainty of Evidence

Currently, AMSTAR 2 is the most widely used tool for evaluating the quality of systematic review methodologies [21]. However, the revision was not intended to deal with the special requirements of diagnostic systematic reviews [21]. Therefore, the methodological quality of the included articles was rated using only the JBI Critical Appraisal Checklist for Systematic Reviews and Research Syntheses [19]. This process was done independently by 2 investigators (MMX and WY), with group discussion when necessary.

The Grading of Recommendations, Assessment, Development, and Evaluation (GRADE) is the established tool for assessing the overall certainty of evidence [21-23]. For this review, we examined the following GRADE domains: (1) risk of bias in the individual studies, (2) inconsistency, (3) indirectness, (4) imprecision, and (5) publication bias. Based on these criteria, the meta-analytical evidence was classified as high, moderate, low, or very low. High and moderate certainty in evidence means that it is very likely or probable that the true effect lies close to the estimated finding, and low or very low certainty means that we have little or very little confidence in the finding [24,25]. This assessment was conducted by 2 independent researchers (QC and QB), with discussion and agreement for any differences.

Narrative Synthesis

Given the significant heterogeneity between the reviews in terms of AI classifiers, imaging modalities, characteristics of participants, settings of conducted studies, etc, a narrative-only summary was considered. Study findings will be presented in graphs in the same way as the syntheses are reported in the narrative text to facilitate the comparison of findings from each included study [26].

Results

Study Selection

The detailed process for this umbrella review is presented in a schematic flowchart in Figure 1. The initial database search retrieved 3690 candidate publications, of which 2691 were screened following the removal of 999 duplicates. Based on the title and abstract, 2350 publications were excluded. The remaining 341 publications underwent a full-text examination to confirm eligibility. This resulted in a final total of 158 studies that met the inclusion criteria (Figure 1). The list of the excluded records during the process of full-text checking is provided in Multimedia Appendix 3.

**Figure 1.** Flow diagram of the study selection process.
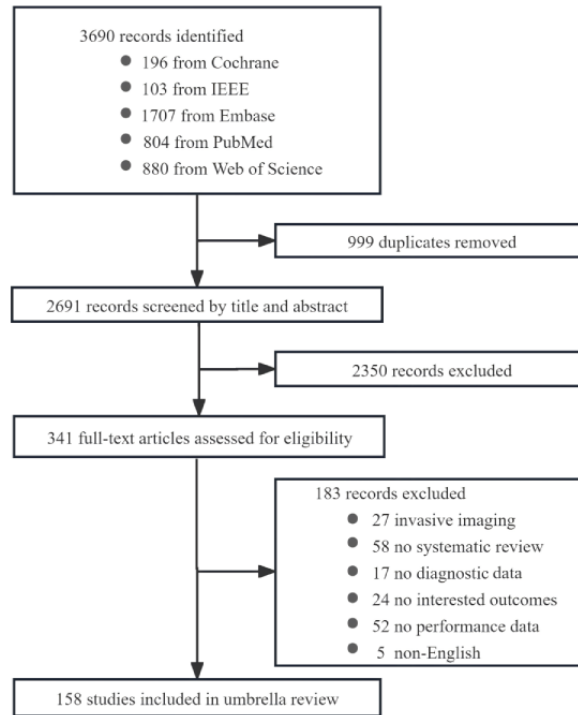
Study Characteristics

The detailed characteristics of the included studies are presented in Figures 2 and 3 and Multimedia Appendices 4-6. More than half of the included studies (133/158, 84.2%) were published in 2021-2024 (Figure 2A). The included studies focused on cancers in the following 8 systems (Figure 2B): central nervous system (CNS) (n=15), head and neck (n=18), respiratory system (n=16), digestive system (n=43), urinary system (n=18), female-related systems (n=23), skin (n=18), and others (n=8). Among them, 1 review reported on both nervous system and respiratory system tumors. The included studies were conducted in 31 different countries, with half of them conducted in Asia (n=86) (Figure 2C). The most commonly searched database was PubMed (113/158, 71.5%) (Figure 2D). The number of retrieved and included studies in the meta-analyses or systematic reviews ranged from 279 to 117,464 and 3 to 32, respectively.

Figure 2. Characteristics of included systematic reviews by (A) year, (B) tumor site, (C) region, and (D) search database. CT: computed tomography; PET: positron emission tomography; MRI: magnetic resonance imaging; US: ultrasound.
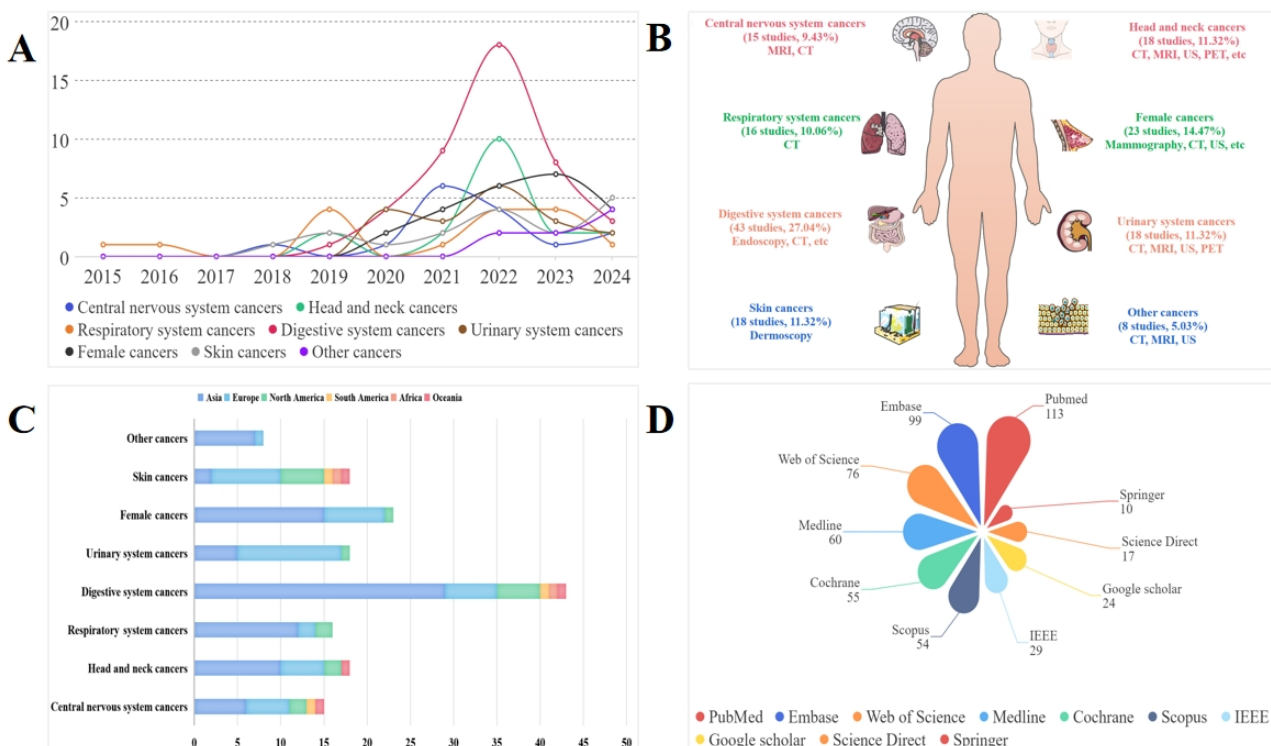
Different proportions of meta-analyses or systematic reviews had registered protocols, followed the PRISMA guidelines, restricted searches to English-language studies, and conducted quality assessment (Figure 3). Several criteria were mentioned for quality assessment, such as Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2), radiomics quality score (RQS), and Checklist for Artificial Intelligence in Medical

Imaging (CLAIM). At least two independent reviewers carried out the study selection process, data extraction, and quality assessment in 106, 82, and 66 meta-analyses or systematic reviews, respectively. Seventy-seven meta-analyses or systematic reviews synthesized the data using quantitative analysis.

Figure 3. Characteristics of included systematic reviews by methodology.
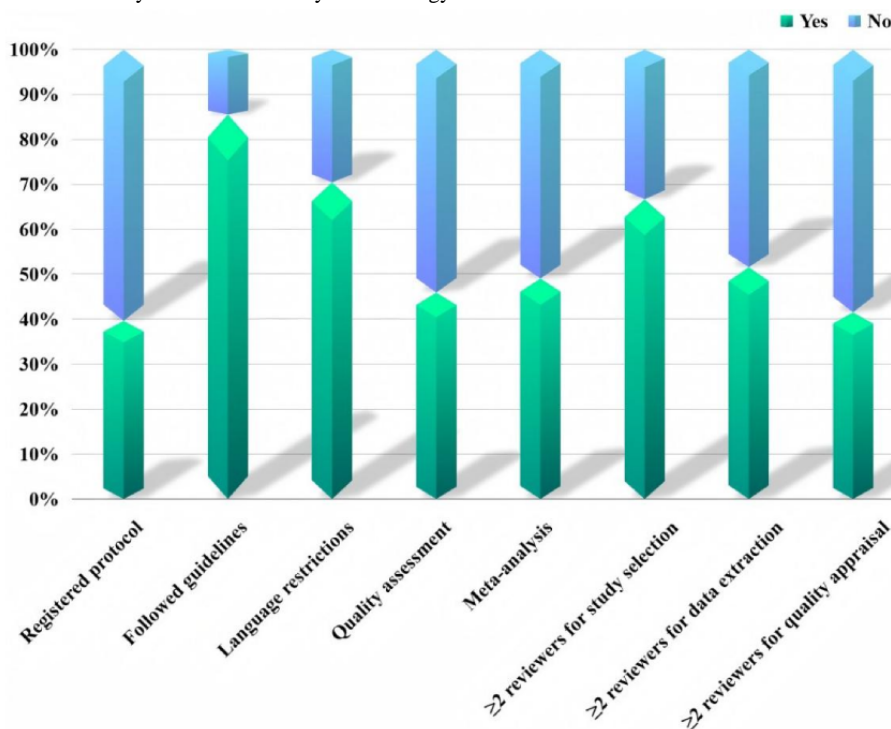
The studies included in the meta-analyses or systematic reviews used various imaging data to match the model. Among them, US, CT, and MRI were the most frequently adopted imaging approaches in most diseases. Moreover, dermoscopy in skin cancer, endoscopy in gastrointestinal tumors, and mammography in breast cancer were employed. A variety of algorithms were used in the included studies, and the most common ones were convolutional neural network (CNN) and support vector machine (SVM). Algorithms, including DenseNet, EfficientNet, RetinaNet, ImageNet, and Inception-v3, were frequently

employed. Forty-nine meta-analyses or systematic reviews included articles that used both internal and external validation.

Quality Appraisal Results

The JBI critical appraisal results are summarized in Figure 4 and Multimedia Appendix 7. Overall, 80.4% (127/158) of the meta-analyses or systematic reviews were of high quality, and the rest were of medium quality. Out of 11 criteria, the minimum number of criteria met was 5 and the maximum was 11. Criteria 1-3, relating to clear questions, appropriate criteria, and appropriate search strategy, were met by the vast majority of included meta-analyses or systematic reviews.

Figure 4. Quality of the included reviews assessed according to the Joanna Briggs Institute (JBI) tool. NA: not applicable.
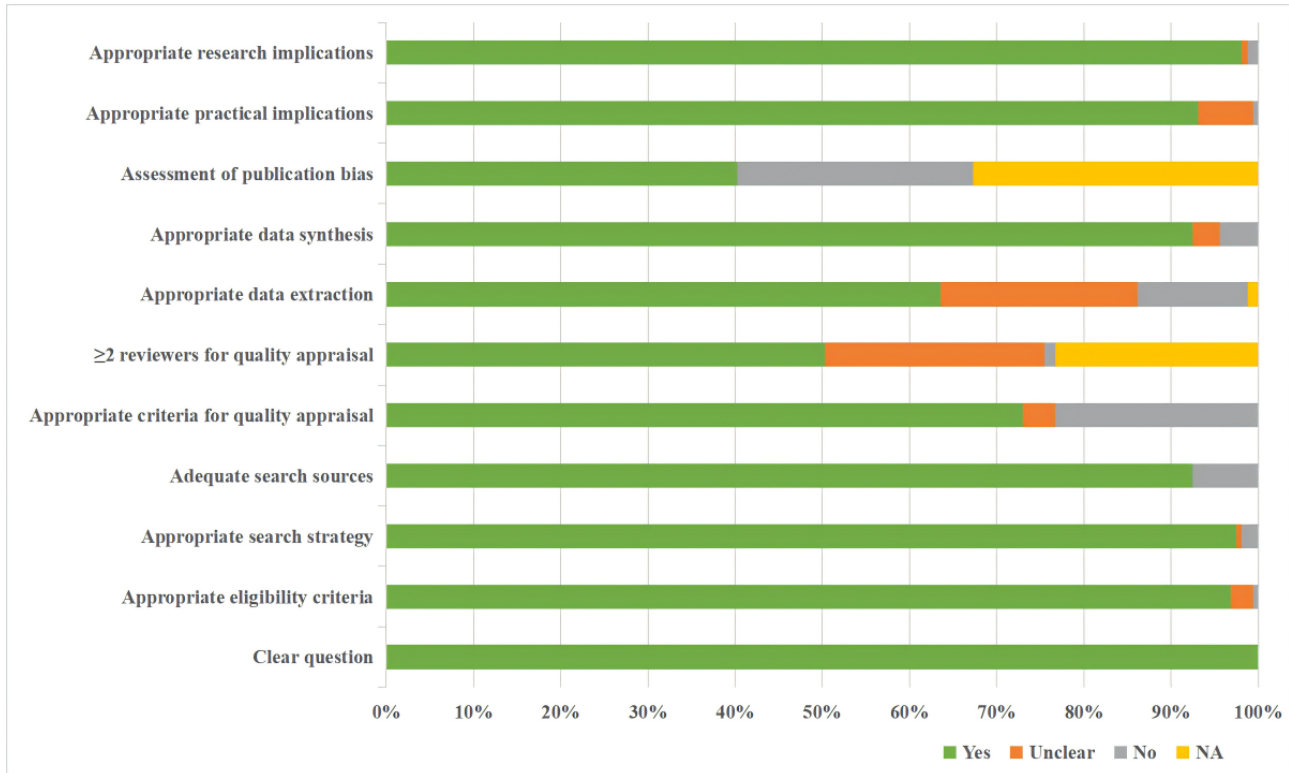

Grading of Evidence

The results of the GRADE assessment are presented in Multimedia Appendix 8 and Figure 5. Overall, 139 pieces of evidence met the conditions of the GRADE. The certainties of evidence were as follows: high, 5.0% (7/139); moderate, 37.4% (52/139); low, 46.8% (65/139); and very low, 10.8% (15/139).

Figure 5. Evidence certainty of the included reviews assessed using the Grading of Recommendations, Assessment, Development, and Evaluations (GRADE) criteria. A: central nervous system cancers; B: head and neck cancers; C: respiratory system cancers; D: digestive system cancers; E: urinary system cancers; F: female cancers; G: skin cancers; H: other cancers.
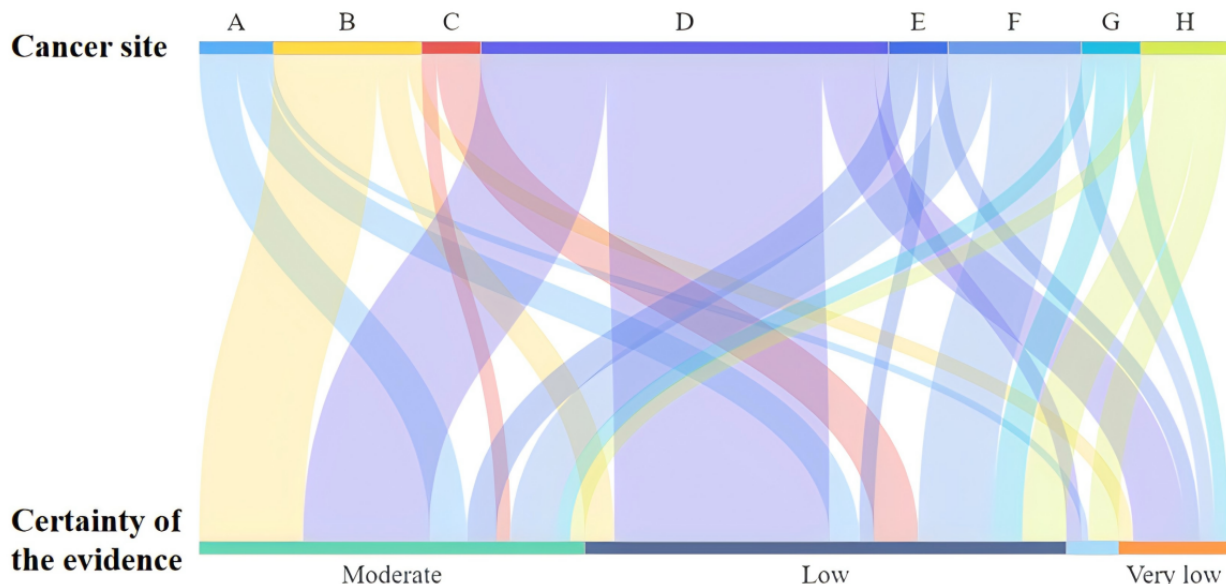
Summary of the Outcome

Overall, 158 meta-analyses or systematic reviews contained information on the diagnostic performance of AI for multisite cancers: brain, oral, nasopharyngeal, thyroid, lung, gastrointestinal, esophageal, liver, pancreas, renal, bladder, prostate, breast, cervical, ovarian, skin, chondrosarcoma, bone, adrenal, and soft tissue tumors. The performance of the AI models in diagnosing these cancers is presented in Multimedia Appendix 9.

CNS Cancers

Eight meta-analyses and 7 systematic reviews [27-41] evaluated the performance of AI algorithms in classifying CNS cancers, mainly using MRI, and 6 of these focused more specifically on the identification of gliomas. The accuracy, sensitivity, and specificity of the classifiers in the reviews ranged from 48% to 100%, 51% to 100%, and 40% to 97%, respectively. Among them, a recently published meta-analysis with a more comprehensive database search showed that the pooled sensitivity, specificity, and AUC were 94% (95% CI 91%-95%), 93% (95% CI 91%-95%), and 0.98 (95% CI 0.96-0.99), respectively [37]. However, the analysis showed an I^2 of 97.6% for sensitivity and I^2 of 96.7% for specificity, and other meta-analyses showed similar heterogeneity. Based on the JBI tool, 93% (14/15) of reviews were of high quality in CNS cancers. According to the GRADE criteria, 50% (5/10) of the evidence was rated as moderate.

Head and Neck Cancers

Eleven meta-analyses and 7 systematic reviews [42-59] summarized the performance of AI in using multiple imaging data to diagnose head and neck cancers, including oral, larynx, pharynx, and thyroid lesions. Several representative systematic reviews deserve to be highlighted. Rokhshad et al [44] reviewed deep learning (DL) applications for detecting head and neck cancer, using radiographic data. The accuracy, sensitivity, and specificity varied from 82.6% to 100%, 74% to 99.7%, and 66.6% to 90.1%, respectively. In a recently published meta-analysis of AI evaluating images of oral lesions, the pooled sensitivity and specificity were 88% (95% CI 87%-88%) and 80% (95% CI 80%-81%), respectively [47]. According to a meta-analysis of the largest number of included studies, the pooled sensitivity and specificity for radiomics in distinguishing thyroid nodules were 87% (95% CI 86%-87%) and 84% (95% CI 84%-85%), respectively [57]. In addition, the estimated sensitivity and specificity of the differential diagnosis between benign and malignant laryngeal lesions were 91% (95% CI 85%-97%) and 94% (95% CI 89%-100%), respectively [45]. With I^2 ranges from 62% to 98.3%, the above pooled analyses revealed a significant variation between studies [44,45,47,48,50-52,54-57]. According to the GRADE criteria, 70% (14/20) of the evidence was judged as moderate.

Respiratory System Cancers

Six meta-analyses and 10 systematic reviews [33,60-74] examined the performance of AI in lung cancer assessment, mainly using CT. The results from 5 meta-analyses indicated that the pooled sensitivity was greater than 84%, while 1 reported a value of 77% [62]. Nevertheless, the pooled specificity was relatively low, mainly distributed around 65% to 80%. It has been suggested that AI-assisted CT diagnostic technology for the classification of pulmonary nodules as benign or malignant has good diagnostic performance, but its specificity needs to be improved. The GRADE criteria indicated 75% (6/8) low-quality evidence, suggesting that caution should be exercised when applying these findings in a clinical setting.

Digestive System Cancers

Forty-three meta-analyses or systematic reviews [75-117] explored the performance of AI in digestive system cancers, including cholangiocarcinoma (n=1), colorectal neoplasia (n=1), esophageal cancer (n=11), gastric cancer (n=19), hepatocellular carcinoma (n=4), and pancreatic cancer (n=7). Among these systematic reviews, 9 meta-analyzed sensitivity, specificity, and AUC and showed ranges of 90% to 95%, 80% to 93.8%, and 0.88 to 0.97, respectively, for esophageal cancer diagnostic performance [77-80,82-86]. For gastric cancer, 2 recently published meta-analyses showed aggregate AUC values greater than 0.94 [95,105]. Additionally, by using QUADAS-2, the included pooled studies were generally rated as high quality. Salehi et al [110] recently published meta-analyses related to hepatocellular carcinoma and reported that the pooled sensitivity and specificity for internally validated AI algorithms were 84% (95% CI 81%-87%) and 92% (95% CI 90%-94%), respectively, and those for externally validated AI algorithms were 85% (95% CI 78%-89%) and 84% (95% CI 72%-91%), respectively. They further found that the diagnostic performance of MRI images was slightly higher than that of CT and US images [110]. Five meta-analyses examined the performance of AI in diagnosing pancreatic cancer, using endoscopic US [111-113,115,116], reporting ranges of 90.4%-93% and 84%-95% for combined sensitivity and specificity, respectively. The evidence quality for the digestive system was classified according to the GRADE criteria as follows: high, 3% (2/55); medium, 31% (17/55); low, 53% (29/55); and very low, 13% (7/55).

Urinary System Cancers

Eighteen meta-analyses or systematic reviews [118-135] examined the performance of AI in urinary system cancers, using CT, PET, and MRI data. In a representative meta-analysis of bladder cancer [120], 8 studies with a total of 860 patients were included. The summary estimates for sensitivity and specificity in predicting bladder cancer were 82% (95% CI 77%-86%) and 81% (95% CI 76%-85%), respectively. More importantly, there was no relevant heterogeneity in diagnostic accuracy measures (I^2=33% and 41% for sensitivity and specificity, respectively). In the systematic reviews of prostate cancer diagnosis [121-134], the accuracy, sensitivity, and specificity of the classifiers ranged from 29.7% to 100%, 22% to 100%, and 6% to 100%, respectively. Among them, a well-structured meta-analysis reported pooled sensitivity and specificity of 81.5% (95% CI 41%-99%) and 83% (95% CI 42%-99%), respectively. The calculated heterogeneity values for the pooled sensitivity and specificity were 84% and 79%, respectively (P<.001) [128]. The GRADE assessment revealed that 50% (4/8) of the evidence was of moderate quality, while the remainder had either low or very low quality.

Female Cancers

Twenty-three studies focused on female systemic tumors [17,136-157]. It is noteworthy that 15 meta-analyses or systematic reviews [136-150] examined the performance of AI in breast cancer detection. Eight of these pooled the results using meta-analysis, and the calculated pooled sensitivity and specificity ranged from 75.4% to 92% and 83% to 90.6%, respectively. It is worth mentioning that in the study by Oh et

al [149], MRI demonstrated a sensitivity of 91% (95% CI 89%-92%) and specificity of 84% (95% CI 82%-86%), mammography-based radiomic features demonstrated a sensitivity of 79% (95% CI 76%-82%) and specificity of 81% (95% CI 79%-84%), and US-based analysis yielded a sensitivity of 92% (95% CI 90%-94%) and specificity of 85% (95% CI 83%-88%). The aforementioned studies exhibited similar diagnostic test accuracy for differentiating benign and malignant breast lesions via radiomic analysis. For ovarian cancer, 4 meta-analyses reported that the ranges of sensitivity and specificity were 75%-94% [17,154-156]. Furthermore, 2 meta-analyses performed subgroup analysis of imaging modalities, with AUC values of 0.94 and 0.95, 0.82 and 0.82, and 0.92 and 0.90 for US, CT, and MRI, respectively [17,154]. The GRADE evaluation indicated that 33% (6/18) of the evidence was of moderate quality. It is notable that this proportion implies a certain level of reliability in the presented findings but also suggests the need for further validation and refinement.

Skin Cancers

Four meta-analyses and 14 systematic reviews [10,158-174] explored the role of AI in skin cancer assessment and exhibited favorable performance. For instance, a newly published meta-analysis [172] found sensitivity and specificity values of 87% and 77.1%, respectively, for AI algorithms, and 79.8% and 73.6%, respectively, for all clinicians (overall). The differences were statistically significant. Moreover, pooled data from 70 studies presented the summary estimates of computer-aided diagnosis systems for melanoma, with a sensitivity of 74% (95% CI 66%-80%) and specificity of 84% (95% CI 79%-88%) [161]. The GRADE evaluation revealed that 75% of the evidence was of either low or very low quality. Admittedly, the included studies presented diverse methodologies and significant heterogeneity regarding the types of images included, the characteristics of the participants, and the methodology for presenting the results.

Other Cancers

Several systematic reviews have focused on some uncommon tumors. For instance, only 1 review evaluated the clinical value of radiomic analysis for chondrosarcoma, and the models for the differential diagnosis of chondrosarcoma, although showing good performance, were considered to have weak evidence [175]. A meta-analysis investigated the performance of CT-based radiomics in diagnosing malignant adrenal tumors [176] and showed an overall pooled AUC of 0.88 (95% CI 0.85-0.91). Dai et al [177] and Zhu et al [178] systematically evaluated the application value of radiomics and DL in the differential diagnosis of benign and malignant soft tissue tumors and demonstrated a pooled sensitivity of 84% and 84% and specificity of 88% and 63%, respectively. Two systematic reviews [179,180] evaluated the diagnostic performance of AI algorithms in the detection of bone malignancy, and the pooled sensitivity and specificity for AI algorithms in 1 meta-analysis were 84% (95% CI 75%-90%) and 91% (95% CI 83%-96%), respectively. In addition, Bai et al [181] reported the diagnostic performance of an AI model for the early detection of lymphoma for the first time, with a pooled sensitivity of 87% (95% CI

83%-91%), specificity of 94% (95% CI 92%-96%), and AUC of 0.97 (95% CI 0.95-0.98).

Discussion

Principal Findings

This umbrella review presents an overview of the AI-image techniques in the domain of cancer diagnosis, using the JBI tool and GRADE guidelines. We included 158 pertinent systematic reviews or meta-analyses involving 8 major human systems, and the GRADE assessment revealed that the overall quality of the evidence was moderate to very low. Although the diagnostic performance of the classifiers varied greatly, most meta-analyses showed positive summary performance. This study witnessed the outstanding performance of AI algorithms in analyzing tumor images, which can accurately identify the characteristics and subtle changes of tumors. However, this emerging field still faces many uncertainties and problems that need to be solved, requiring in-depth discussions with scientific and objective attitudes.

Representativity of Study Objects

Ideally, the diagnostic approach should be explored in the setting of its intended use. In reality, numerous studies were not representative of the general population and community practice, with studies being performed in an experimental environment. For example, to date, all published data investigating radiomics in the field of pancreatic neuroendocrine tumors were mainly single-centered and retrospective studies [182]. The datasets largely included lesions from patients recruited in specialist clinical settings, and the applicability of these classifiers remains theoretical, making it more difficult to replicate the findings [183]. Meanwhile, the number of positive images and negative images in some included studies was significantly different, which was highlighted in a review on gastrointestinal cancer [92]. As suggested in recently published guidelines [88,184], a step forward needs to be achieved. Both current and future studies need to be constantly refined and externally validated in multicentric, large-sample, randomized clinical trials. In this regard, public access to large databases of clinical and radiological correlated data will be instrumental [182]. The creation of large public image datasets with images as representative as possible of the world's people to avoid racial bias is a major task in this research field [171,185].

Trends and Challenges of AI Technology

The variety of algorithms used in the assessed studies was high, and among them, CNN and SVM were the most widely used DL and machine learning methods in classification steps. SVM was typically used for limited data, while CNN was preferred for larger datasets. This distinction reflects the researcher's expertise and preferences but is primarily driven by data constraints [186]. Even if high-performance models generate reasonable results, data from different training environments can contain various types of bias and noise and may not be reproduced again [159,187], and technique challenges remain to be settled [170]. The ability of clinicians to accept and trust the outputs of an algorithm, when the decision-making process is not apparent or comprehensible to them, may prove to be an

obstacle to adoption [188]. A vivid metaphor is that the algorithm is mostly like a "black box" [189]. Acceptability to medical professionals and regulatory agencies may be increased if there is enhanced understanding as to how an algorithm arrives at its decision. Moving forward, further research into so called "explainable AI" [190] may provide the necessary transparency, trust, and accountability desired by the health care profession.

It should be discussed that a wide range of metrics was employed to report diagnostic performance in AI studies. We noticed that the evaluations were not performed based on a single procedure, and many studies used their own indicators. For example, some studies reported only 1 indicator; hence, it is not possible to know whether the study performed better in terms of other indicators [186]. There is no uniform standard and evaluation standard for the specific AI model design. In view of certain parameters being comprehensive and useful in clinical practice, it is an ideal research process to compare the accuracy, sensitivity, specificity, or AUC of some models with those of other methods and datasets in studies [100].

In the domain of tumor imaging diagnosis with the application of AI, several notable trends have emerged. Regarding AI technology trends, the continuous refinement and enhancement of DL algorithms, especially the advancement of CNN, have significantly improved the precision in analyzing features from tumor images. For instance, advanced architectures, such as ResNet and DenseNet, have achieved remarkable success in enhancing diagnostic accuracy [37,51,52]. Moreover, the use of transfer learning and pretrained models has become increasingly prevalent. By leveraging pretrained models on large-scale general image datasets like ImageNet and fine-tuning them for tumor imaging data, training time and data requirements have been considerably reduced [43]. In terms of target task categories, tumor detection has emerged as a critical objective, where AI techniques strive to identify early tumor indications from subtle image variations [87,114]. The classification and grading of tumors have also received significant attention, helping physicians in formulating more individualized treatment plans [191]. Additionally, in the process of literature screening, we found that comparing pre- and posttreatment imaging to determine whether the tumor has shrunk or remained stable, has become an important area of focus [192,193].

The emergence of generative AI and transformer-based technology has brought new possibilities. Generative AI can generate realistic tumor simulation images, providing a novel approach for expanding training data and aiding physicians in better comprehending tumor morphological changes [194]. Transformer-based technology excels in handling sequential data, such as time series in medical imaging, contributing to capturing the dynamic development of tumors [195,196]. When it comes to foundation models, the exploration of large-scale ones, such as large language models, in the medical field is on the rise. Integrating image and related clinical text information enables more comprehensive diagnostic support [197-199]. Multi-modal fusion foundation models are emerging, integrating information from multiple imaging modalities like CT, MRI, and PET to enhance diagnostic accuracy and reliability [200]. As mentioned above, the trends in AI for tumor imaging

diagnosis are moving toward greater precision, comprehensiveness, personalization, and efficiency, offering new prospects for the detection and further precise treatment of tumors.

AI, Clinicians, and Patients

According to the findings of this study, the AI systems in cancer detection follow 2 directions: (1) as concurrent assistants aiding clinicians in diagnosis and (2) as standalone systems providing an independent assessment. In the former case, the AI system is used as an aid during interpretation [201]. In particular, when using images for diagnosis, factors, such as physician fatigue, stress, and limited experience, may lead to neglected or misdiagnosed diseases. In contrast, AI can continuously provide reliable performance over a short period of time, has the potential to compensate for limited human capabilities, can prevent mistakes made by doctors in clinical practice, and can facilitate the training and education of less experienced endoscopists [116]. In the latter case, AI can play an important role in areas where the economy is underdeveloped and medical resources are relatively scarce. However, a major concern is the reproducibility, reliability, and usability of the diagnostic capability of AI, as false-negative results will lead to misdiagnosis and false-positive ones may result in overdiagnosis [159,202]. In this respect, AI may not completely replace doctors, and human beings and machines working together in harmony is an ideal state that can result in optimum performance.

As accuracy becomes more important, consolidating the information found in various forms of clinical data is vital in future clinical practice [96]. If purely based on image modeling, it is evident that the approach does not take into account all the information a clinician would rely on to evaluate a difficult examination [203]. The addition of clinical data, such as race, age, and gender, as inputs for classifiers may help to increase classification accuracy. This supplemental data could benefit clinicians in their decision-making [171]. Furthermore, AI has shown its advantages in computational power and learning capacity. An AI model is adept at integrating a lot of information from most data, which has the potential to reduce the workload of clinicians substantially, as it is difficult for them to integrate complex data manually [96]. These aspects should be included in future work.

Multidisciplinary Cooperation

Processing and modeling medical images have traditionally represented complex tasks requiring multidisciplinary collaboration. Technically, the lack of a suitable information technology infrastructure in medical practices, the absence of high-quality curated data, and difficulties to access and exchange data are inhibiting the translation of integrated diagnostics into clinical routine and even research. Looking at the issue from a legal perspective, the sensitive issues of data privacy and security, patient consent, and autonomy must be fully considered. This means that from a legal perspective, data acquisition, storage, transfer, processing, and analysis will have to comply with all laws, regulations, and further legal requirements. In addition, the law and its interpretation and implementation must constantly adapt to the evolving

state-of-the-art in technology [204,205]. We have observed a growing trend of clinical studies using commercial software for image processing [12]. The cost-effectiveness evaluation of AI in diagnosis is also warranted. Multidirectional interactions should be implemented among physicians, developers, legislators, and economists to foster multidisciplinary collaboration to tackle these challenges jointly [71]. When malpractice cases involving medical AI applications arise, multidisciplinary teams need to provide necessary suggestions [190].

Quality Assessment Tools of the Included Studies

Multiple tools were employed in our included systematic reviews to comprehensively assess the quality of the included articles. The studies were evaluated by the RQS [206], CLAIM [207], Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD) statement [208], Image Biomarker Standardization Initiative (IBSI) guideline [209], and QUADAS-2/AI tool [116,210]. It should be noted that each tool has its own focus. For example, the RQS does not take into account sample size or how the model actually performs but instead represents an evaluation of both how rigorous model development is and how impactful the study may be to the field [211]. The IBSI was able to produce and validate reference values for radiomics features. The TRIPOD is a similar example that aims to promote the transparent reporting of diagnostic accuracy model studies. There are suggestions from researchers that CLAIM may guide the update of the TRIPOD and RQS, because it not only includes general reporting criteria but also allows extra distinction of unique shortness in DL. Considering the overlapping items and high correlation between these tools, researchers can choose an appropriate one or a combination as the case may be. We are also confident that relevant organizations or scholars will continue to refine existing tools to better guide and evaluate relevant research.

Strengths and Limitations

The strengths of this umbrella review include an exhaustive search of medical and engineering databases; assessment of methodological quality using the JBI tool; independent screening, quality assessment, and data extraction processes; and methodological rigor in the conduct of the review achieved by following the PRISMA guidelines. We also identified challenges that will need to be overcome for the technology to be implemented into daily clinical practice [94]. Umbrella reviews represent the highest level of evidence synthesis currently available and are becoming increasingly influential in biomedical literature [212]. Evidence from the umbrella review highlights the massive opportunity for AI image–based cancer detection. Remarkably, by identifying potential biases and methodological limitations, the information provided can raise the awareness of health care decision makers regarding the strengths and weaknesses of available AI diagnostic tools.

Several limitations of this review merit consideration. First, this review included a synthesis of evidence from existing systematic reviews or meta-analyses [116]. Therefore, cancers that were not systematically reviewed in the pre-existing literature were not included in this umbrella review. The lack of generalizability and the problem of out-of-distribution lesions, such as rare or more harmful disease categories, are limitations that are not addressed by current studies. Second, we could not examine the algorithms that were built without publication, such as those developed for private companies or for business usage [12]. Third, this umbrella review was limited by substantial variability in terms of the characteristics of the patients included, method quality, and study design, which precluded a full statistical analysis. Furthermore, the GRADE assessment revealed that the overall quality of nearly half of the evidence was low (47%). Since secondary research is highly dependent on the quality and availability of primary studies, both this umbrella review and previous systematic reviews were considerably restricted in improving the current understanding of cancer diagnosis.

Conclusion

Conclusively, the main points discussed and investigated in this review can be summarized as follows. AI algorithms have shown remarkable efficacy in the noninvasive imaging diagnosis of tumors. Notwithstanding the progress, certain challenges persist. The data used in current AI models often have limitations in terms of quantity, quality, and diversity, which can compromise the model's generalization ability. Additionally, the lack of transparency and interpretability of some AI systems poses difficulties for clinicians in understanding and trusting the diagnostic results. In the future, efforts should be directed toward enhancing data quality and diversity, developing more interpretable AI models, and conducting large-scale, multicenter clinical trials to validate the practical effectiveness of these technologies. While AI holds tremendous promise in noninvasive tumor imaging diagnosis, a concerted effort from the research community, clinicians, and policymakers is required to overcome the existing hurdles and translate this potential into improved patient outcomes and health care delivery.

Authors' Contributions

HLX, TTG, XJS, and QJW contributed to the study design. HLX, TTG, XJS, QB, WY, and MMX contributed to data collection and analysis. HLX, TTG, CL, MG, YS, HZS, SG, YHZ, and QJW wrote the first draft of the manuscript and edited the manuscript. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) checklist.
[[XLSX File (Microsoft Excel File), 14 KB](#)-[Multimedia Appendix 1](#)]

Multimedia Appendix 2

Search terms and search strategy.
[[XLSX File (Microsoft Excel File), 11 KB](#)-[Multimedia Appendix 2](#)]

Multimedia Appendix 3

List of the excluded records during the process of full-text checking.
[[XLSX File (Microsoft Excel File), 32 KB](#)-[Multimedia Appendix 3](#)]

Multimedia Appendix 4

Metadata of the included reviews.
[[XLSX File (Microsoft Excel File), 51 KB](#)-[Multimedia Appendix 4](#)]

Multimedia Appendix 5

Search sources, search results, study selection, data extraction, quality assessment, and data synthesis in the included reviews.
[[XLSX File (Microsoft Excel File), 49 KB](#)-[Multimedia Appendix 5](#)]

Multimedia Appendix 6

Features of models in the included reviews.
[[XLSX File (Microsoft Excel File), 53 KB](#)-[Multimedia Appendix 6](#)]

Multimedia Appendix 7

Reviewer judgements about each quality criterion for each included review.
[[XLSX File (Microsoft Excel File), 51 KB](#)-[Multimedia Appendix 7](#)]

Multimedia Appendix 8

The results of the Grading of Recommendations, Assessment, Development, and Evaluations (GRADE) assessment of the evidence certainty of diagnostic accuracy regarding the association between artificial intelligence image systems and multisite cancers.
[[XLSX File (Microsoft Excel File), 27 KB](#)-[Multimedia Appendix 8](#)]

Multimedia Appendix 9

Classifier performance in differentiating multiple system cancers.
[[XLSX File (Microsoft Excel File), 52 KB](#)-[Multimedia Appendix 9](#)]

References

1. Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2024;74(3):229-263. [[FREE Full text](#)] [doi: [10.3322/caac.21834](#)] [Medline: [38572751](#)]
2. Siegel RL, Giaquinto AN, Jemal A. Cancer statistics, 2024. CA Cancer J Clin. 2024;74(1):12-49. [[FREE Full text](#)] [doi: [10.3322/caac.21820](#)] [Medline: [38230766](#)]
3. Zheng T, Pierre-Pierre N, Yan X, Huo Q, Almodovar AJO, Valerio F, et al. Gold nanoparticle-enabled blood test for early stage cancer detection and risk assessment. ACS Appl Mater Interfaces. Apr 01, 2015;7(12):6819-6827. [doi: [10.1021/acsami.5b00371](#)] [Medline: [25757512](#)]
4. Meng X, Pang X, Zhang K, Gong C, Yang J, Dong H, et al. Recent advances in near-infrared-II fluorescence imaging for deep-tissue molecular analysis and cancer diagnosis. Small. Aug 2022;18(31):e2202035. [doi: [10.1002/smll.202202035](#)] [Medline: [35762403](#)]

5. Kurmi Y, Chaurasia V, Ganesh N. Tumor malignancy detection using histopathology imaging. J Med Imaging Radiat Sci. Dec 2019;50(4):514-528. [doi: 10.1016/j.jmir.2019.07.004] [Medline: 31501064]

6. He L, Long LR, Antani S, Thoma GR. Histology image analysis for carcinoma detection and grading. Comput Methods Programs Biomed. Sep 2012;107(3):538-556. [FREE Full text] [doi: 10.1016/j.cmpb.2011.12.007] [Medline: 22436890]

7. Hung C. Computational algorithms on medical image processing. Curr Med Imaging. 2020;16(5):467-468. [doi: 10.2174/1573405616052004101447443] [Medline: 32484080]

8. Soomro TA, Zheng L, Afifi AJ, Ali A, Soomro S, Yin M, et al. Image segmentation for MR brain tumor detection using machine learning: a review. IEEE Rev Biomed Eng. 2023;16:70-90. [doi: 10.1109/RBME.2022.3185292] [Medline: 35737636]

9. Bi WL, Hosny A, Schabath MB, Giger ML, Birkbak NJ, Mehrtash A, et al. Artificial intelligence in cancer imaging: Clinical challenges and applications. CA Cancer J Clin. Mar 2019;69(2):127-157. [FREE Full text] [doi: 10.3322/caac.21552] [Medline: 30720861]

10. Marka A, Carter JB, Toto E, Hassanpour S. Automated detection of nonmelanoma skin cancer using digital images: a systematic review. BMC Med Imaging. Feb 28, 2019;19(1):21. [FREE Full text] [doi: 10.1186/s12880-019-0307-7] [Medline: 30819133]

11. Rajpurkar P, Lungren MP. The current and future state of AI interpretation of medical images. N Engl J Med. May 25, 2023;388(21):1981-1990. [doi: 10.1056/NEJMra2301725] [Medline: 37224199]

12. van Leeuwen KG, Schalekamp S, Rutten MJCM, van Ginneken B, de Rooij M. Artificial intelligence in radiology: 100 commercially available products and their scientific evidence. Eur Radiol. Jun 2021;31(6):3797-3804. [FREE Full text] [doi: 10.1007/s00330-021-07892-z] [Medline: 33856519]

13. Hollon T, Jiang C, Chowdury A, Nasir-Moin M, Kondepudi A, Aabedi A, et al. Artificial-intelligence-based molecular classification of diffuse gliomas using rapid, label-free optical imaging. Nat Med. Apr 2023;29(4):828-832. [FREE Full text] [doi: 10.1038/s41591-023-02252-4] [Medline: 36959422]

14. Bhinder B, Gilvary C, Madhukar NS, Elemento O. Artificial intelligence in cancer research and precision medicine. Cancer Discov. Apr 2021;11(4):900-915. [FREE Full text] [doi: 10.1158/2159-8290.CD-21-0090] [Medline: 33811123]

15. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. Dermatologist-level classification of skin cancer with deep neural networks. Nature. Feb 02, 2017;542(7639):115-118. [FREE Full text] [doi: 10.1038/nature21056] [Medline: 28117445]

16. Rodriguez-Ruiz A, Lång K, Gubern-Merida A, Broeders M, Gennaro G, Clauser P, et al. Stand-alone artificial intelligence for breast cancer detection in mammography: comparison with 101 radiologists. J Natl Cancer Inst. Sep 01, 2019;111(9):916-922. [FREE Full text] [doi: 10.1093/jnci/djy222] [Medline: 30834436]

17. Xu H, Gong T, Liu F, Chen H, Xiao Q, Hou Y, et al. Artificial intelligence performance in image-based ovarian cancer identification: A systematic review and meta-analysis. EClinicalMedicine. Nov 2022;53:101662. [FREE Full text] [doi: 10.1016/j.eclinm.2022.101662] [Medline: 36147628]

18. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. Mar 29, 2021;372:n71. [FREE Full text] [doi: 10.1136/bmj.n71] [Medline: 33782057]

19. Aromataris E, Fernandez R, Godfrey CM, Holly C, Khalil H, Tungpunkom P. Summarizing systematic reviews: methodological development, conduct and reporting of an umbrella review approach. Int J Evid Based Healthc. Sep 2015;13(3):132-140. [doi: 10.1097/XEB.0000000000000055] [Medline: 26360830]

20. Wang J, Wu W, Chang K, Chen L, Chi S, Kara M, et al. Ultrasound imaging for the diagnosis and evaluation of sarcopenia: an umbrella review. Life (Basel). Dec 22, 2021;12(1):9. [FREE Full text] [doi: 10.3390/life12010009] [Medline: 35054402]

21. Shea BJ, Reeves BC, Wells G, Thuku M, Hamel C, Moran J, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. BMJ. Sep 21, 2017;358:j4008. [FREE Full text] [doi: 10.1136/bmj.j4008] [Medline: 28935701]

22. Horvath AR. Grading quality of evidence and strength of recommendations for diagnostic tests and strategies. Clin Chem. May 2009;55(5):853-855. [doi: 10.1373/clinchem.2008.117614] [Medline: 19299540]

23. Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, et al. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. J Clin Epidemiol. Apr 2011;64(4):383-394. [doi: 10.1016/j.jclinepi.2010.04.026] [Medline: 21195583]

24. Patikorn C, Roubal K, Veettil SK, Chandran V, Pham T, Lee YY, et al. Intermittent fasting and obesity-related health outcomes: an umbrella review of meta-analyses of randomized clinical trials. JAMA Netw Open. Dec 01, 2021;4(12):e2139558. [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.39558] [Medline: 34919135]

25. Chiavaroli L, Viguiliouk E, Nishi SK, Blanco Mejia S, Rahelić D, Kahleová H, et al. DASH dietary pattern and cardiometabolic outcomes: an umbrella review of systematic reviews and meta-analyses. Nutrients. Feb 05, 2019;11(2):338. [FREE Full text] [doi: 10.3390/nu11020338] [Medline: 30764511]

26. Abd-Alrazaq A, Alhuwail D, Schneider J, Toro CT, Ahmed A, Alzubaidi M, et al. The performance of artificial intelligence-driven technologies in diagnosing mental disorders: an umbrella review. NPJ Digit Med. Jul 07, 2022;5(1):87. [FREE Full text] [doi: 10.1038/s41746-022-00631-8] [Medline: 35798934]

27. Al-Galal SAY, Alshaikhli IFT, Abdulrazzaq MM. MRI brain tumor medical images analysis using deep learning techniques: a systematic review. Health Technol. Jan 14, 2021;11(2):267-282. [doi: 10.1007/s12553-020-00514-6]

28. Naeem A, Anees T, Naqvi RA, Loh W. A comprehensive analysis of recent deep and federated-learning-based methodologies for brain tumor diagnosis. J Pers Med. Feb 13, 2022;12(2):275. [FREE Full text] [doi: 10.3390/jpm12020275] [Medline: 35207763]

29. Kouli O, Hassane A, Badran D, Kouli T, Hossain-Ibrahim K, Steele JD. Automated brain tumor identification using magnetic resonance imaging: A systematic review and meta-analysis. Neurooncol Adv. 2022;4(1):vdac081. [FREE Full text] [doi: 10.1093/noajnl/vdac081] [Medline: 35769411]

30. Xie Y, Zaccagna F, Rundo L, Testa C, Agati R, Lodi R, et al. Convolutional neural network techniques for brain tumor classification (from 2015 to 2022): review, challenges, and future perspectives. Diagnostics (Basel). Jul 31, 2022;12(8):1850. [FREE Full text] [doi: 10.3390/diagnostics12081850] [Medline: 36010200]

31. Badrigilan S, Nabavi S, Abin AA, Rostampour N, Abedi I, Shirvani A, et al. Deep learning approaches for automated classification and segmentation of head and neck cancers and brain tumors in magnetic resonance images: a meta-analysis study. Int J Comput Assist Radiol Surg. Apr 2021;16(4):529-542. [doi: 10.1007/s11548-021-02326-z] [Medline: 33666859]

32. Silva Santana L, Borges Camargo Diniz J, Mothé Glioche Gasparri L, Buccaran Canto A, Batista Dos Reis S, Santana Neville Ribeiro I, et al. Application of machine learning for classification of brain tumors: a systematic review and meta-analysis. World Neurosurg. Jun 2024;186:204-218.e2. [doi: 10.1016/j.wneu.2024.03.152] [Medline: 38580093]

33. Tandon R, Agrawal S, Rathore NPS, Mishra AK, Jain SK. A systematic review on deep learning-based automated cancer diagnosis models. J Cell Mol Med. Mar 2024;28(6):e18144. [FREE Full text] [doi: 10.1111/jcmm.18144] [Medline: 38426930]

34. Buchlak QD, Esmaili N, Leveque J, Bennett C, Farrokhi F, Piccardi M. Machine learning applications to neuroimaging for glioma detection and classification: An artificial intelligence augmented systematic review. J Clin Neurosci. Jul 2021;89:177-198. [doi: 10.1016/j.jocn.2021.04.043] [Medline: 34119265]

35. van Kempen EJ, Post M, Mannil M, Witkam RL, Ter Laan M, Patel A, et al. Performance of machine learning algorithms for glioma segmentation of brain MRI: a systematic literature review and meta-analysis. Eur Radiol. Dec 2021;31(12):9638-9653. [FREE Full text] [doi: 10.1007/s00330-021-08035-0] [Medline: 34019128]

36. Sohn CK, Bisdas S. Diagnostic accuracy of machine learning-based radiomics in grading gliomas: systematic review and meta-analysis. Contrast Media Mol Imaging. 2020;2020:2127062. [FREE Full text] [doi: 10.1155/2020/2127062] [Medline: 33746649]

37. Sun W, Song C, Tang C, Pan C, Xue P, Fan J, et al. Performance of deep learning algorithms to distinguish high-grade glioma from low-grade glioma: A systematic review and meta-analysis. iScience. Jun 16, 2023;26(6):106815. [FREE Full text] [doi: 10.1016/j.isci.2023.106815] [Medline: 37250800]

38. Alhasan AS. Clinical applications of artificial intelligence, machine learning, and deep learning in the imaging of gliomas: a systematic review. Cureus. Nov 2021;13(11):e19580. [FREE Full text] [doi: 10.7759/cureus.19580] [Medline: 34926051]

39. van Kempen EJ, Post M, Mannil M, Kusters B, Ter Laan M, Meijer FJA, et al. Accuracy of machine learning algorithms for the classification of molecular features of gliomas on MRI: a systematic literature review and meta-analysis. Cancers (Basel). May 26, 2021;13(11):2606. [FREE Full text] [doi: 10.3390/cancers13112606] [Medline: 34073309]

40. Nguyen AV, Blears EE, Ross E, Lall RR, Ortega-Barnett J. Machine learning applications for the differentiation of primary central nervous system lymphoma from glioblastoma on imaging: a systematic review and meta-analysis. Neurosurg Focus. Nov 01, 2018;45(5):E5. [doi: 10.3171/2018.8.FOCUS18325] [Medline: 30453459]

41. Cassinelli Petersen GI, Shatalov J, Verma T, Brim WR, Subramanian H, Brackett A, et al. Machine learning in differentiating gliomas from primary CNS lymphomas: a systematic review, reporting quality, and risk of bias assessment. AJNR Am J Neuroradiol. Apr 2022;43(4):526-533. [FREE Full text] [doi: 10.3174/ajnr.A7473] [Medline: 35361577]

42. DeJohn CR, Grant SR, Seshadri M. Application of machine learning methods to improve the performance of ultrasound in head and neck oncology: a literature review. Cancers (Basel). Jan 28, 2022;14(3):665. [FREE Full text] [doi: 10.3390/cancers14030665] [Medline: 35158932]

43. Hasan Z, Key S, Habib A, Wong E, Aweidah L, Kumar A, et al. Convolutional neural networks in ENT radiology: systematic review of the literature. Ann Otol Rhinol Laryngol. Apr 2023;132(4):417-430. [doi: 10.1177/00034894221095899] [Medline: 35651308]

44. Rokhshad R, Salehi SN, Yavari A, Shobeiri P, Esmaeili M, Manila N, et al. Deep learning for diagnosis of head and neck cancers through radiographic data: a systematic review and meta-analysis. Oral Radiol. Jan 2024;40(1):1-20. [doi: 10.1007/s11282-023-00715-5] [Medline: 37855976]

45. Żurek M, Jasak K, Niemczyk K, Rzepakowska A. Artificial intelligence in laryngeal endoscopy: systematic review and meta-analysis. J Clin Med. May 12, 2022;11(10):2752. [FREE Full text] [doi: 10.3390/jcm11102752] [Medline: 35628878]

46. Ng WT, But B, Choi HCW, de Bree R, Lee AWM, Lee VHF, et al. Application of artificial intelligence for nasopharyngeal carcinoma management - a systematic review. Cancer Manag Res. 2022;14:339-366. [FREE Full text] [doi: 10.2147/CMAR.S341583] [Medline: 35115832]

XSL•FO

RenderX

47. Li C, Chen X, Chen C, Gong Z, Pataer P, Liu X, et al. Application of deep learning radiomics in oral squamous cell carcinoma-Extracting more information from medical images using advanced feature analysis. J Stomatol Oral Maxillofac Surg. Jun 2024;125(3S):101840. [doi: 10.1016/j.jormas.2024.101840] [Medline: 38548062]

48. Elmakaty I, Elmarasi M, Amarah A, Abdo R, Malki MI. Accuracy of artificial intelligence-assisted detection of Oral Squamous Cell Carcinoma: A systematic review and meta-analysis. Crit Rev Oncol Hematol. Oct 2022;178:103777. [doi: 10.1016/j.critrevonc.2022.103777] [Medline: 35931404]

49. de Chauveron J, Unger M, Lescaille G, Wendling L, Kurtz C, Rochefort J. Artificial intelligence for oral squamous cell carcinoma detection based on oral photographs: A comprehensive literature review. Cancer Med. Jan 2024;13(1):e6822. [FREE Full text] [doi: 10.1002/cam4.6822] [Medline: 38164652]

50. Han R, Lin N, Huang J, Ma X. Diagnostic accuracy of Raman spectroscopy in oral squamous cell carcinoma. Front Oncol. 2022;12:925032. [FREE Full text] [doi: 10.3389/fonc.2022.925032] [Medline: 35992884]

51. Kim J, Kim BG, Hwang SH. Efficacy of artificial intelligence-assisted discrimination of oral cancerous lesions from normal mucosa based on the oral mucosal image: a systematic review and meta-analysis. Cancers (Basel). Jul 19, 2022;14(14):3499. [FREE Full text] [doi: 10.3390/cancers14143499] [Medline: 35884560]

52. Ferro A, Kotecha S, Fan K. Machine learning in point-of-care automated classification of oral potentially malignant and malignant disorders: a systematic review and meta-analysis. Sci Rep. Aug 13, 2022;12(1):13797. [FREE Full text] [doi: 10.1038/s41598-022-17489-1] [Medline: 35963880]

53. Khanagar SB, Naik S, Al Kheraif AA, Vishwanathaiah S, Maganur PC, Alhazmi Y, et al. Application and performance of artificial intelligence technology in oral cancer diagnosis and prediction of prognosis: a systematic review. Diagnostics (Basel). May 31, 2021;11(6):1004. [FREE Full text] [doi: 10.3390/diagnostics11061004] [Medline: 34072804]

54. Potipimpanon P, Charakorn N, Hirunwiwatkul P. A comparison of artificial intelligence versus radiologists in the diagnosis of thyroid nodules using ultrasonography: a systematic review and meta-analysis. Eur Arch Otorhinolaryngol. Nov 2022;279(11):5363-5373. [doi: 10.1007/s00405-022-07436-1] [Medline: 35767056]

55. Zhao W, Fu L, Huang Z, Zhu J, Ma B. Effectiveness evaluation of computer-aided diagnosis system for the diagnosis of thyroid nodules on ultrasound: A systematic review and meta-analysis. Medicine (Baltimore). Aug 2019;98(32):e16379. [FREE Full text] [doi: 10.1097/MD.0000000000016379] [Medline: 31393347]

56. Xu L, Gao J, Wang Q, Yin J, Yu P, Bai B, et al. Computer-aided diagnosis systems in diagnosing malignant thyroid nodules on ultrasonography: a systematic review and meta-analysis. Eur Thyroid J. Jul 2020;9(4):186-193. [FREE Full text] [doi: 10.1159/000504390] [Medline: 32903956]

57. Cleere EF, Davey MG, O'Neill S, Corbett M, O'Donnell JP, Hacking S, et al. Radiomic detection of malignancy within thyroid nodules using ultrasonography-a systematic review and meta-analysis. Diagnostics (Basel). Mar 24, 2022;12(4):794. [FREE Full text] [doi: 10.3390/diagnostics12040794] [Medline: 35453841]

58. Sharifi-Zahabi E, Sharafabad FH, Abdollahzad H, Malekahmadi M, Rad NB. Circulating advanced glycation end products and their soluble receptors in relation to all-cause and cardiovascular mortality: a systematic review and meta-analysis of prospective observational studies. Adv Nutr. Dec 01, 2021;12(6):2157-2171. [FREE Full text] [doi: 10.1093/advances/nmab072] [Medline: 34139010]

59. Conn Busch JM, Cozzi JL, Li H, Lan L, Giger ML, Keutgen XM. Role of machine learning in differentiating benign from malignant indeterminate thyroid nodules: A literature review. Health Sciences Review. Jun 2023;7:100089. [doi: 10.1016/j.hsr.2023.100089]

60. Thong LT, Chou HS, Chew HSJ, Lau Y. Diagnostic test accuracy of artificial intelligence-based imaging for lung cancer screening: A systematic review and meta-analysis. Lung Cancer. Feb 2023;176:4-13. [doi: 10.1016/j.lungcan.2022.12.002] [Medline: 36566582]

61. Liu M, Wu J, Wang N, Zhang X, Bai Y, Guo J, et al. The value of artificial intelligence in the diagnosis of lung cancer: A systematic review and meta-analysis. PLoS One. 2023;18(3):e0273445. [FREE Full text] [doi: 10.1371/journal.pone.0273445] [Medline: 36952523]

62. Sokouti M, Sokouti M, Sokouti B. A systematic review and meta-analysis on performance of intelligent systems in lung cancer: Where are we? Artif Intell Rev. Sep 18, 2019;53(5):3287-3298. [doi: 10.1007/s10462-019-09764-x]

63. Pacurari AC, Bhattarai S, Muhammad A, Avram C, Mederle AO, Rosca O, et al. Diagnostic accuracy of machine learning ai architectures in detection and classification of lung cancer: a systematic review. Diagnostics (Basel). Jun 22, 2023;13(13):2145. [FREE Full text] [doi: 10.3390/diagnostics13132145] [Medline: 37443539]

64. Valente IRS, Cortez PC, Neto EC, Soares JM, de Albuquerque VHC, Tavares JMR. Automatic 3D pulmonary nodule detection in CT images: A survey. Comput Methods Programs Biomed. Feb 2016;124:91-107. [FREE Full text] [doi: 10.1016/j.cmpb.2015.10.006] [Medline: 26652979]

65. Amir GJ, Lehmann HP. After detection: the improved accuracy of lung cancer assessment using radiologic computer-aided diagnosis. Acad Radiol. Feb 2016;23(2):186-191. [FREE Full text] [doi: 10.1016/j.acra.2015.10.014] [Medline: 26616209]

66. Zhang G, Yang Z, Gong L, Jiang S, Wang L, Cao X, et al. An appraisal of nodule diagnosis for lung cancer in CT images. J Med Syst. May 15, 2019;43(7):181. [doi: 10.1007/s10916-019-1327-0] [Medline: 31093830]

67. Jin H, Yu C, Gong Z, Zheng R, Zhao Y, Fu Q. Machine learning techniques for pulmonary nodule computer-aided diagnosis using CT images: A systematic review. Biomedical Signal Processing and Control. Jan 2023;79:104104. [doi: 10.1016/j.bspc.2022.104104]

68. Wang X, Mao K, Wang L, Yang P, Lu D, He P. An appraisal of lung nodules automatic classification algorithms for CT images. Sensors (Basel). Jan 07, 2019;19(1):194. [FREE Full text] [doi: 10.3390/s19010194] [Medline: 30621101]

69. Li D, Mikela Vilmun B, Frederik Carlsen J, Albrecht-Beste E, Ammitzbøl Lauridsen C, Bachmann Nielsen M, et al. The performance of deep learning algorithms on automatic pulmonary nodule detection and classification tested on different datasets that are not derived from LIDC-IDRI: a systematic review. Diagnostics (Basel). Nov 29, 2019;9(4):207. [FREE Full text] [doi: 10.3390/diagnostics9040207] [Medline: 31795409]

70. Huang G, Wei X, Tang H, Bai F, Lin X, Xue D. A systematic review and meta-analysis of diagnostic performance and physicians' perceptions of artificial intelligence (AI)-assisted CT diagnostic technology for the classification of pulmonary nodules. J Thorac Dis. Aug 2021;13(8):4797-4811. [FREE Full text] [doi: 10.21037/jtd-21-810] [Medline: 34527320]

71. Min Y, Hu L, Wei L, Nie S. Computer-aided detection of pulmonary nodules based on convolutional neural networks: a review. Phys Med Biol. Mar 07, 2022;67(6):ac568e. [doi: 10.1088/1361-6560/ac568e] [Medline: 35180709]

72. Forte GC, Altmayer S, Silva RF, Stefani MT, Libermann LL, Cavion CC, et al. Deep learning algorithms for diagnosis of lung cancer: a systematic review and meta-analysis. Cancers (Basel). Aug 09, 2022;14(16):3856. [FREE Full text] [doi: 10.3390/cancers14163856] [Medline: 36010850]

73. Jassim M, Jaber M. ystematic review for lung cancer detection and lung nodule classification: Taxonomy, challenges, and recommendation future works. Journal of Intelligent Systems. 2022;31(1):964. [FREE Full text] [doi: 10.1515/jisys-2022-0062]

74. Zheng X, He B, Hu Y, Ren M, Chen Z, Zhang Z, et al. Diagnostic accuracy of deep learning and radiomics in lung cancer staging: a systematic review and meta-analysis. Front Public Health. 2022;10:938113. [FREE Full text] [doi: 10.3389/fpubh.2022.938113] [Medline: 35923964]

75. Njei B, McCarty TR, Mohan BP, Fozo L, Navaneethan U. Artificial intelligence in endoscopic imaging for detection of malignant biliary strictures and cholangiocarcinoma: a systematic review. Ann Gastroenterol. 2023;36(2):223-230. [FREE Full text] [doi: 10.20524/aog.2023.0779] [Medline: 36864938]

76. Yamada M, Saito Y, Yamada S, Kondo H, Hamamoto R. Detection of flat colorectal neoplasia by artificial intelligence: A systematic review. Best Pract Res Clin Gastroenterol. 2021;52-53:101745. [doi: 10.1016/j.bpg.2021.101745] [Medline: 34172250]

77. Ma H, Wang L, Chen Y, Tian L. Convolutional neural network-based artificial intelligence for the diagnosis of early esophageal cancer based on endoscopic images: A meta-analysis. Saudi J Gastroenterol. 2022;28(5):332-340. [FREE Full text] [doi: 10.4103/sjg.sjg_178_22] [Medline: 35848703]

78. Zhang SM, Wang YJ, Zhang ST. Accuracy of artificial intelligence-assisted detection of esophageal cancer and neoplasms on endoscopic images: A systematic review and meta-analysis. J Dig Dis. Jun 2021;22(6):318-328. [FREE Full text] [doi: 10.1111/1751-2980.12992] [Medline: 33871932]

79. Bang CS, Lee JJ, Baik GH. Computer-aided diagnosis of esophageal cancer and neoplasms in endoscopic images: a systematic review and meta-analysis of diagnostic test accuracy. Gastrointest Endosc. May 2021;93(5):1006-1015.e13. [FREE Full text] [doi: 10.1016/j.gie.2020.11.025] [Medline: 33290771]

80. Bhatti KM, Khanzada ZS, Kuzman M, Ali SM, Iftikhar SY, Small P. Diagnostic performance of artificial intelligence-based models for the detection of early esophageal cancers in Barret's esophagus: a meta-analysis of patient-based studies. Cureus. Jun 2021;13(6):e15447. [FREE Full text] [doi: 10.7759/cureus.15447] [Medline: 34258114]

81. Syed T, Doshi A, Guleria S, Syed S, Shah T. Artificial intelligence and its role in identifying esophageal neoplasia. Dig Dis Sci. Dec 2020;65(12):3448-3455. [FREE Full text] [doi: 10.1007/s10620-020-06643-2] [Medline: 33057945]

82. Zhang J, Mi J, Wang R. Application of convolutional neural network-based endoscopic imaging in esophageal cancer or high-grade dysplasia: A systematic review and meta-analysis. World J Gastrointest Oncol. Nov 15, 2023;15(11):1998-2016. [FREE Full text] [doi: 10.4251/wjgo.v15.i11.1998] [Medline: 38077641]

83. Guidozzi N, Menon N, Chidambaram S, Markar SR. The role of artificial intelligence in the endoscopic diagnosis of esophageal cancer: a systematic review and meta-analysis. Dis Esophagus. Nov 30, 2023;36(12):doad048. [FREE Full text] [doi: 10.1093/dote/doad048] [Medline: 37480192]

84. Islam MM, Poly TN, Walther BA, Yeh C, Seyed-Abdul S, Li YJ, et al. Deep learning for the diagnosis of esophageal cancer in endoscopic images: a systematic review and meta-analysis. Cancers (Basel). Dec 05, 2022;14(23):5996. [FREE Full text] [doi: 10.3390/cancers14235996] [Medline: 36497480]

85. Visaggi P, Barberio B, Gregori D, Azzolina D, Martinato M, Hassan C, et al. Systematic review with meta-analysis: artificial intelligence in the diagnosis of oesophageal diseases. Aliment Pharmacol Ther. Mar 2022;55(5):528-540. [FREE Full text] [doi: 10.1111/apt.16778] [Medline: 35098562]

86. Tan JL, Chinnaratha MA, Woodman R, Martin R, Chen H, Carneiro G, et al. Diagnostic accuracy of artificial intelligence (AI) to detect early neoplasia in Barrett's esophagus: a non-comparative systematic review and meta-analysis. Front Med (Lausanne). 2022;9:890720. [FREE Full text] [doi: 10.3389/fmed.2022.890720] [Medline: 35814747]

87. Hosseini F, Asadi F, Emami H, Ebnali M. Machine learning applications for early detection of esophageal cancer: a systematic review. BMC Med Inform Decis Mak. Jul 17, 2023;23(1):124. [FREE Full text] [doi: 10.1186/s12911-023-02235-y] [Medline: 37460991]

88. Lui TKL, Tsui VWM, Leung WK. Accuracy of artificial intelligence-assisted detection of upper GI lesions: a systematic review and meta-analysis. Gastrointest Endosc. Oct 2020;92(4):821-830.e9. [doi: 10.1016/j.gie.2020.06.034] [Medline: 32562608]

89. Mohan BP, Khan SR, Kassab LL, Ponnada S, Dulai PS, Kochhar GS. Accuracy of convolutional neural network-based artificial intelligence in diagnosis of gastrointestinal lesions based on endoscopic images: A systematic review and meta-analysis. Endosc Int Open. Nov 2020;8(11):E1584-E1594. [FREE Full text] [doi: 10.1055/a-1236-3007] [Medline: 33140014]

90. Yan T, Wong PK, Qin Y. Deep learning for diagnosis of precancerous lesions in upper gastrointestinal endoscopy: A review. World J Gastroenterol. May 28, 2021;27(20):2531-2544. [FREE Full text] [doi: 10.3748/wjg.v27.i20.2531] [Medline: 34092974]

91. Arribas J, Antonelli G, Frazzoni L, Fuccio L, Ebigbo A, van der Sommen F, et al. Standalone performance of artificial intelligence for upper GI neoplasia: a meta-analysis. Gut. Oct 30, 2020:321922. [doi: 10.1136/gutjnl-2020-321922] [Medline: 33127833]

92. Luo D, Kuang F, Du J, Zhou M, Liu X, Luo X, et al. Artificial intelligence-assisted endoscopic diagnosis of early upper gastrointestinal cancer: a systematic review and meta-analysis. Front Oncol. 2022;12:855175. [FREE Full text] [doi: 10.3389/fonc.2022.855175] [Medline: 35756602]

93. Langarizadeh M, Sayadi M. Machine learning techniques for diagnosis of lower gastrointestinal cancer: a systematic review. Iranian Red Crescent Medical Journal. 2021;23(7):7. [FREE Full text] [doi: 10.32592/ircmj.2021.23.7.436]

94. Chidambaram S, Sounderajah V, Maynard N, Markar SR. Diagnostic performance of artificial intelligence-centred systems in the diagnosis and postoperative surveillance of upper gastrointestinal malignancies using computed tomography imaging: a systematic review and meta-analysis of diagnostic accuracy. Ann Surg Oncol. Mar 2022;29(3):1977-1990. [FREE Full text] [doi: 10.1245/s10434-021-10882-6] [Medline: 34762214]

95. Shi Y, Fan H, Li L, Hou Y, Qian F, Zhuang M, et al. The value of machine learning approaches in the diagnosis of early gastric cancer: a systematic review and meta-analysis. World J Surg Oncol. Feb 01, 2024;22(1):40. [FREE Full text] [doi: 10.1186/s12957-024-03321-9] [Medline: 38297303]

96. Bhardwaj P, Bhandari G, Kumar Y, Gupta S. An investigational approach for the prediction of gastric cancer using artificial intelligence techniques: a systematic review. Arch Computat Methods Eng. Apr 19, 2022;29(6):4379-4400. [doi: 10.1007/s11831-022-09737-4]

97. Jin P, Ji X, Kang W, Li Y, Liu H, Ma F, et al. Artificial intelligence in gastric cancer: a systematic review. J Cancer Res Clin Oncol. Sep 2020;146(9):2339-2350. [doi: 10.1007/s00432-020-03304-9] [Medline: 32613386]

98. Islam MM, Poly TN, Walther BA, Lin M, Li YJ. Artificial intelligence in gastric cancer: identifying gastric cancer using endoscopic images with convolutional neural network. Cancers (Basel). Oct 20, 2021;13(21):5253. [FREE Full text] [doi: 10.3390/cancers13215253] [Medline: 34771416]

99. Xie F, Zhang K, Li F, Ma G, Ni Y, Zhang W, et al. Diagnostic accuracy of convolutional neural network-based endoscopic image analysis in diagnosing gastric cancer and predicting its invasion depth: a systematic review and meta-analysis. Gastrointest Endosc. Apr 2022;95(4):599-609.e7. [doi: 10.1016/j.gie.2021.12.021] [Medline: 34979114]

100. Zhao Y, Hu B, Wang Y, Yin X, Jiang Y, Zhu X. Identification of gastric cancer with convolutional neural networks: a systematic review. Multimed Tools Appl. 2022;81(8):11717-11736. [FREE Full text] [doi: 10.1007/s11042-022-12258-8] [Medline: 35221775]

101. Chen P, Lu Y, Kang Y, Chang C. The accuracy of artificial intelligence in the endoscopic diagnosis of early gastric cancer: pooled analysis study. J Med Internet Res. May 16, 2022;24(5):e27694. [FREE Full text] [doi: 10.2196/27694] [Medline: 35576561]

102. Jiang K, Jiang X, Pan J, Wen Y, Huang Y, Weng S, et al. Current evidence and future perspective of accuracy of artificial intelligence application for early gastric cancer diagnosis with endoscopy: a systematic and meta-analysis. Front Med (Lausanne). 2021;8:629080. [FREE Full text] [doi: 10.3389/fmed.2021.629080] [Medline: 33791323]

103. Dilaghi E, Lahner E, Annibale B, Esposito G. Systematic review and meta-analysis: Artificial intelligence for the diagnosis of gastric precancerous lesions and Helicobacter pylori infection. Dig Liver Dis. Dec 2022;54(12):1630-1638. [doi: 10.1016/j.dld.2022.03.007] [Medline: 35382973]

104. Ye XH, Zhao LL, Wang L. Diagnostic accuracy of endoscopic ultrasound with artificial intelligence for gastrointestinal stromal tumors: A meta-analysis. J Dig Dis. May 2022;23(5-6):253-261. [doi: 10.1111/1751-2980.13110] [Medline: 35793389]

105. Gomes RSA, de Oliveira GHP, de Moura DTH, Kotinda APST, Matsubayashi CO, Hirsch BS, et al. Endoscopic ultrasound artificial intelligence-assisted for prediction of gastrointestinal stromal tumors diagnosis: A systematic review and meta-analysis. World J Gastrointest Endosc. Aug 16, 2023;15(8):528-539. [FREE Full text] [doi: 10.4253/wjge.v15.i8.528] [Medline: 37663113]

106. Liu X, Song W, Mao T, Zhang Q, Zhang C, Li X. Application of artificial intelligence in the diagnosis of subepithelial lesions using endoscopic ultrasonography: a systematic review and meta-analysis. Front Oncol. 2022;12:915481. [FREE Full text] [doi: 10.3389/fonc.2022.915481] [Medline: 36046054]

107. Azer SA. Deep learning with convolutional neural networks for identification of liver masses and hepatocellular carcinoma: A systematic review. World J Gastrointest Oncol. Dec 15, 2019;11(12):1218-1230. [FREE Full text] [doi: 10.4251/wjgo.v11.i12.1218] [Medline: 31908726]

108. Wei Q, Tan N, Xiong S, Luo W, Xia H, Luo B. Deep learning methods in medical image-based hepatocellular carcinoma diagnosis: a systematic review and meta-analysis. Cancers (Basel). Dec 03, 2023;15(23):5701. [FREE Full text] [doi: 10.3390/cancers15235701] [Medline: 38067404]

109. Bo Z, Song J, He Q, Chen B, Chen Z, Xie X, et al. Application of artificial intelligence radiomics in the diagnosis, treatment, and prognosis of hepatocellular carcinoma. Comput Biol Med. May 2024;173:108337. [FREE Full text] [doi: 10.1016/j.compbiomed.2024.108337] [Medline: 38547656]

110. Salehi MA, Harandi H, Mohammadi S, Shahrabi Farahani M, Shojaei S, Saleh RR. Diagnostic performance of artificial intelligence in detection of hepatocellular carcinoma: a meta-analysis. J Imaging Inform Med. Aug 2024;37(4):1297-1311. [doi: 10.1007/s10278-024-01058-1] [Medline: 38438694]

111. Prasoppokakorn T, Tiyarattanachai T, Chaiteerakij R, Decharatanachart P, Mekaroonkamol P, Ridtitid W, et al. Application of artificial intelligence for diagnosis of pancreatic ductal adenocarcinoma by EUS: A systematic review and meta-analysis. Endosc Ultrasound. 2022;11(1):17-26. [FREE Full text] [doi: 10.4103/EUS-D-20-00219] [Medline: 34937308]

112. Dumitrescu EA, Ungureanu BS, Cazacu IM, Florescu LM, Streba L, Croitoru VM, et al. Diagnostic value of artificial intelligence-assisted endoscopic ultrasound for pancreatic cancer: a systematic review and meta-analysis. Diagnostics (Basel). Jan 25, 2022;12(2):309. [FREE Full text] [doi: 10.3390/diagnostics12020309] [Medline: 35204400]

113. Mohan BP, Facciorusso A, Khan SR, Madhu D, Kassab LL, Ponnada S, et al. Pooled diagnostic parameters of artificial intelligence in EUS image analysis of the pancreas: A descriptive quantitative review. Endosc Ultrasound. 2022;11(3):156-169. [FREE Full text] [doi: 10.4103/EUS-D-21-00063] [Medline: 35313417]

114. Jan Z, El Assadi F, Abd-Alrazaq A, Jithesh PV. Artificial intelligence for the prediction and early diagnosis of pancreatic cancer: scoping review. J Med Internet Res. Mar 31, 2023;25:e44248. [FREE Full text] [doi: 10.2196/44248] [Medline: 37000507]

115. Lv B, Wang K, Wei N, Yu F, Tao T, Shi Y. Diagnostic value of deep learning-assisted endoscopic ultrasound for pancreatic tumors: a systematic review and meta-analysis. Front Oncol. 2023;13:1191008. [FREE Full text] [doi: 10.3389/fonc.2023.1191008] [Medline: 37576885]

116. Yin H, Yang X, Sun L, Pan P, Peng L, Li K, et al. The value of artificial intelligence techniques in predicting pancreatic ductal adenocarcinoma with EUS images: A meta-analysis and systematic review. Endosc Ultrasound. 2023;12(1):50-58. [FREE Full text] [doi: 10.4103/EUS-D-21-00131] [Medline: 35313419]

117. Goyal H, Sherazi SAA, Gupta S, Perisetti A, Achebe I, Ali A, et al. Application of artificial intelligence in diagnosis of pancreatic malignancies by endoscopic ultrasound: a systemic review. Therap Adv Gastroenterol. 2022;15:17562848221093873. [FREE Full text] [doi: 10.1177/17562848221093873] [Medline: 35509425]

118. Salem H, Soria D, Lund JN, Awwad A. A systematic review of the applications of Expert Systems (ES) and machine learning (ML) in clinical urology. BMC Med Inform Decis Mak. Jul 22, 2021;21(1):223. [FREE Full text] [doi: 10.1186/s12911-021-01585-9] [Medline: 34294092]

119. Huang X, Wang X, Lan X, Deng J, Lei Y, Lin F. The role of radiomics with machine learning in the prediction of muscle-invasive bladder cancer: A mini review. Front Oncol. 2022;12:990176. [FREE Full text] [doi: 10.3389/fonc.2022.990176] [Medline: 36059618]

120. Kozikowski M, Suarez-Ibarrola R, Osiecki R, Bilski K, Gratzke C, Shariat SF, et al. Role of radiomics in the prediction of muscle-invasive bladder cancer: a systematic review and meta-analysis. Eur Urol Focus. May 2022;8(3):728-738. [doi: 10.1016/j.euf.2021.05.005] [Medline: 34099417]

121. Cuocolo R, Cipullo MB, Stanzione A, Romeo V, Green R, Cantoni V, et al. Machine learning for the identification of clinically significant prostate cancer on MRI: a meta-analysis. Eur Radiol. Dec 2020;30(12):6877-6887. [doi: 10.1007/s00330-020-07027-w] [Medline: 32607629]

122. Sihotang RC, Agustino C, Huang F, Parikesit D, Rahman F, Hamid ARAH. Accuracy of machine learning models using ultrasound images in prostate cancer diagnosis: a systematic review. Med J Indones. Oct 20, 2023;32(2):112-121. [doi: 10.13181/mji.oa.236765]

123. Sushentsev N, Moreira Da Silva N, Yeung M, Barrett T, Sala E, Roberts M, et al. Comparative performance of fully-automated and semi-automated artificial intelligence methods for the detection of clinically significant prostate cancer on MRI: a systematic review. Insights Imaging. Mar 28, 2022;13(1):59. [FREE Full text] [doi: 10.1186/s13244-022-01199-3] [Medline: 35347462]

124. Michaely HJ, Aringhieri G, Cioni D, Neri E. Current value of biparametric prostate MRI with machine-learning or deep-learning in the detection, grading, and characterization of prostate cancer: a systematic review. Diagnostics (Basel). Mar 24, 2022;12(4):799. [FREE Full text] [doi: 10.3390/diagnostics12040799] [Medline: 35453847]

125. Syer T, Mehta P, Antonelli M, Mallett S, Atkinson D, Ourselin S, et al. Artificial intelligence compared to radiologists for the initial diagnosis of prostate cancer on magnetic resonance imaging: a systematic review and recommendations for future studies. Cancers (Basel). Jul 01, 2021;13(13):3318. [FREE Full text] [doi: 10.3390/cancers13133318] [Medline: 34282762]

126. Arif M, Niessen WJ, Schoots IG, Veenland JF. Automated classification of significant prostate cancer on MRI: a systematic review on the performance of machine learning applications. Cancers (Basel). Jun 17, 2020;12(6):1606. [FREE Full text] [doi: 10.3390/cancers12061606] [Medline: 32560558]

127. Sugano D, Sanford D, Abreu A, Duddalwar V, Gill I, Cacciamani GE. Impact of radiomics on prostate cancer detection: a systematic review of clinical applications. Curr Opin Urol. Nov 2020;30(6):754-781. [doi: 10.1097/MOU.0000000000000822] [Medline: 32941257]

128. Castaldo R, Cavaliere C, Soricelli A, Salvatore M, Pecchia L, Franzese M. Radiomic and genomic machine learning method performance for prostate cancer diagnosis: systematic literature review. J Med Internet Res. Apr 01, 2021;23(4):e22394. [FREE Full text] [doi: 10.2196/22394] [Medline: 33792552]

129. Roest C, Fransen SJ, Kwee TC, Yakar D. Comparative performance of deep learning and radiologists for the diagnosis and localization of clinically significant prostate cancer at MRI: a systematic review. Life (Basel). Sep 26, 2022;12(10):1490. [FREE Full text] [doi: 10.3390/life12101490] [Medline: 36294928]

130. Rouvière O, Jaouen T, Baseilhac P, Benomar ML, Escande R, Crouzet S, et al. Artificial intelligence algorithms aimed at characterizing or detecting prostate cancer on MRI: How accurate are they when tested on independent cohorts? - A systematic review. Diagn Interv Imaging. May 2023;104(5):221-234. [FREE Full text] [doi: 10.1016/j.diii.2022.11.005] [Medline: 36517398]

131. Sihotang RC, Agustino C, Huang F, Parikesit D, Rahman F, Hamid ARAH. The accuracy of machine learning models using ultrasound images in prostate cancer diagnosis: A systematic review. medRxiv. 2022. URL: https://www.medrxiv.org/content/10.1101/2022.02.03.22270377v1 [accessed 2025-03-06]

132. Lindgren Belal S, Frantz S, Minarik D, Enqvist O, Wikström E, Edenbrandt L, et al. Applications of artificial intelligence in PSMA PET/CT for prostate cancer imaging. Semin Nucl Med. Jan 2024;54(1):141-149. [FREE Full text] [doi: 10.1053/j.semnuclmed.2023.06.001] [Medline: 37357026]

133. Talyshinskii A, Hameed BMZ, Ravinder PP, Naik N, Randhawa P, Shah M, et al. Catalyzing precision medicine: artificial intelligence advancements in prostate cancer diagnosis and management. Cancers (Basel). May 09, 2024;16(10):1809. [FREE Full text] [doi: 10.3390/cancers16101809] [Medline: 38791888]

134. Thenault R, Kaulanjan K, Darde T, Rioux-Leclercq N, Bensalah K, Mermier M, et al. The application of artificial intelligence in prostate cancer management—what improvements can be expected? a systematic review. Applied Sciences. Sep 15, 2020;10(18):6428. [doi: 10.3390/app10186428]

135. Song H, Wang X, Wu R, Liu W. The influence of manual segmentation strategies and different phases selection on machine learning-based computed tomography in renal tumors: a systematic review and meta-analysis. Radiol Med. Jul 2024;129(7):1025-1037. [doi: 10.1007/s11547-024-01825-8] [Medline: 38740709]

136. Gardezi SJS, Elazab A, Lei B, Wang T. Breast cancer detection and diagnosis using mammographic data: systematic review. J Med Internet Res. Jul 26, 2019;21(7):e14464. [FREE Full text] [doi: 10.2196/14464] [Medline: 31350843]

137. Li J, Sang T, Yu W, Jiang M, Hunag S, Cao C, et al. The value of S-Detect for the differential diagnosis of breast masses on ultrasound: a systematic review and pooled meta-analysis. Med Ultrason. May 11, 2020;22(2):211-219. [FREE Full text] [doi: 10.11152/mu-2402] [Medline: 32399528]

138. Hanis TM, Islam MA, Musa KI. Diagnostic accuracy of machine learning models on mammography in breast cancer classification: a meta-analysis. Diagnostics (Basel). Jul 05, 2022;12(7):1643. [FREE Full text] [doi: 10.3390/diagnostics12071643] [Medline: 35885548]

139. Wang X, Meng S. Diagnostic accuracy of S-Detect to breast cancer on ultrasonography: A meta-analysis (PRISMA). Medicine (Baltimore). Aug 26, 2022;101(34):e30359. [FREE Full text] [doi: 10.1097/MD.0000000000030359] [Medline: 36042648]

140. Li Z, Ye J, Du H, Cao Y, Wang Y, Liu D, et al. Preoperative prediction power of radiomics for breast cancer: a systemic review and meta-analysis. Front Oncol. 2022;12:837257. [FREE Full text] [doi: 10.3389/fonc.2022.837257] [Medline: 35299744]

141. Omega Boro L, Nandi G. Diagnostic performance of deep learning in screened mammogram: systematic review. Curr Med Imaging. Apr 17, 2023:e170423215871. [FREE Full text] [doi: 10.2174/1573405620666230417085624] [Medline: 37070450]

142. Liu J, Lei J, Ou Y, Zhao Y, Tuo X, Zhang B, et al. Mammography diagnosis of breast cancer screening through machine learning: a systematic review and meta-analysis. Clin Exp Med. Oct 2023;23(6):2341-2356. [doi: 10.1007/s10238-022-00895-0] [Medline: 36242643]

143. Uzun Ozsahin D, Ikechukwu Emegano D, Uzun B, Ozsahin I. The systematic review of artificial intelligence applications in breast cancer diagnosis. Diagnostics (Basel). Dec 23, 2022;13(1):45. [FREE Full text] [doi: 10.3390/diagnostics13010045] [Medline: 36611337]

144. Hickman SE, Woitek R, Le EPV, Im YR, Mouritsen Luxhøj C, Aviles-Rivero AI, et al. Machine learning for workflow applications in screening mammography: systematic review and meta-analysis. Radiology. Jan 2022;302(1):88-104. [FREE Full text] [doi: 10.1148/radiol.2021210391] [Medline: 34665034]

145. Adam R, Dell'Aquila K, Hodges L, Maldjian T, Duong TQ. Deep learning applications to breast cancer detection by magnetic resonance imaging: a literature review. Breast Cancer Res. Jul 24, 2023;25(1):87. [FREE Full text] [doi: 10.1186/s13058-023-01687-4] [Medline: 37488621]

146. Dan Q, Xu Z, Burrows H, Bissram J, Stringer JSA, Li Y. Diagnostic performance of deep learning in ultrasound diagnosis of breast cancer: a systematic review. NPJ Precis Oncol. Jan 27, 2024;8(1):21. [FREE Full text] [doi: 10.1038/s41698-024-00514-z] [Medline: 38280946]

147. Freeman K, Geppert J, Stinton C, Todkill D, Johnson S, Clarke A, et al. Use of artificial intelligence for image analysis in breast cancer screening programmes: systematic review of test accuracy. BMJ. Sep 01, 2021;374:n1872. [FREE Full text] [doi: 10.1136/bmj.n1872] [Medline: 34470740]

148. Aggarwal R, Sounderajah V, Martin G, Ting DSW, Karthikesalingam A, King D, et al. Diagnostic accuracy of deep learning in medical imaging: a systematic review and meta-analysis. NPJ Digit Med. Apr 07, 2021;4(1):65. [FREE Full text] [doi: 10.1038/s41746-021-00438-z] [Medline: 33828217]

149. Oh K, Vasandani N, Anwar A. Radiomics to differentiate malignant and benign breast lesions: a systematic review and diagnostic test accuracy meta-analysis. Cureus. Nov 2023;15(11):e49015. [FREE Full text] [doi: 10.7759/cureus.49015] [Medline: 38024014]

150. Yoon JH, Strand F, Baltzer PAT, Conant EF, Gilbert FJ, Lehman CD, et al. Standalone AI for breast cancer detection at screening digital mammography and digital breast tomosynthesis: a systematic review and meta-analysis. Radiology. Jun 2023;307(5):e222639. [FREE Full text] [doi: 10.1148/radiol.222639] [Medline: 37219445]

151. Viñals R, Jonnalagedda M, Petignat P, Thiran J, Vassilakos P. Artificial intelligence-based cervical cancer screening on images taken during visual inspection with acetic acid: a systematic review. Diagnostics (Basel). Feb 22, 2023;13(5):836. [FREE Full text] [doi: 10.3390/diagnostics13050836] [Medline: 36899979]

152. Taddese AA, Tilahun BC, Awoke T, Atnafu A, Mamuye A, Mengiste SA. Deep-learning models for image-based gynecological cancer diagnosis: a systematic review and meta- analysis. Front Oncol. 2023;13:1216326. [FREE Full text] [doi: 10.3389/fonc.2023.1216326] [Medline: 38273847]

153. Qi L, Li X, Yang Y, Zhao M, Lin A, Ma L. Accuracy of machine learning in the preoperative identification of ovarian borderline tumors: a meta-analysis. Clin Radiol. Jul 2024;79(7):501-514. [doi: 10.1016/j.crad.2024.02.012] [Medline: 38670918]

154. Ma L, Huang L, Chen Y, Zhang L, Nie D, He W, et al. AI diagnostic performance based on multiple imaging modalities for ovarian tumor: A systematic review and meta-analysis. Front Oncol. 2023;13:1133491. [FREE Full text] [doi: 10.3389/fonc.2023.1133491] [Medline: 37152032]

155. Mitchell S, Nikolopoulos M, El-Zarka A, Al-Karawi D, Al-Zaidi S, Ghai A, et al. Artificial intelligence in ultrasound diagnoses of ovarian cancer: a systematic review and meta-analysis. Cancers (Basel). Jan 19, 2024;16(2):422. [FREE Full text] [doi: 10.3390/cancers16020422] [Medline: 38275863]

156. Koch AH, Jeelof LS, Muntinga CLP, Gootzen TA, van de Kruis NMA, Nederend J, et al. Analysis of computer-aided diagnostics in the preoperative diagnosis of ovarian cancer: a systematic review. Insights Imaging. Feb 15, 2023;14(1):34. [FREE Full text] [doi: 10.1186/s13244-022-01345-x] [Medline: 36790570]

157. Akazawa M, Hashimoto K. Artificial intelligence in gynecologic cancers: Current status and future challenges - A systematic review. Artif Intell Med. Oct 2021;120:102164. [doi: 10.1016/j.artmed.2021.102164] [Medline: 34629152]

158. Widaatalla Y, Wolswijk T, Adan F, Hillen LM, Woodruff HC, Halilaj I, et al. The application of artificial intelligence in the detection of basal cell carcinoma: A systematic review. J Eur Acad Dermatol Venereol. Jun 2023;37(6):1160-1167. [doi: 10.1111/jdv.18963] [Medline: 36785993]

159. Zhang S, Wang Y, Zheng Q, Li J, Huang J, Long X. Artificial intelligence in melanoma: A systematic review. J Cosmet Dermatol. Nov 2022;21(11):5993-6004. [doi: 10.1111/jocd.15323] [Medline: 36001057]

160. Popescu D, El-Khatib M, El-Khatib H, Ichim L. New trends in melanoma detection using neural networks: a systematic review. Sensors (Basel). Jan 10, 2022;22(2):496. [FREE Full text] [doi: 10.3390/s22020496] [Medline: 35062458]

161. Dick V, Sinz C, Mittlböck M, Kittler H, Tschandl P. Accuracy of computer-aided diagnosis of melanoma: a meta-analysis. JAMA Dermatol. Nov 01, 2019;155(11):1291-1299. [FREE Full text] [doi: 10.1001/jamadermatol.2019.1375] [Medline: 31215969]

162. Rajpara SM, Botello AP, Townend J, Ormerod AD. Systematic review of dermoscopy and digital dermoscopy/ artificial intelligence for the diagnosis of melanoma. Br J Dermatol. Sep 2009;161(3):591-604. [doi: 10.1111/j.1365-2133.2009.09093.x] [Medline: 19302072]

163. Jeong HK, Park C, Henao R, Kheterpal M. Deep learning in dermatology: a systematic review of current approaches, outcomes, and limitations. JID Innov. Jan 2023;3(1):100150. [FREE Full text] [doi: 10.1016/j.xjidi.2022.100150] [Medline: 36655135]

164. Miller I, Rosic N, Stapelberg M, Hudson J, Coxon P, Furness J, et al. Performance of commercial dermatoscopic systems that incorporate artificial intelligence for the identification of melanoma in general practice: a systematic review. Cancers (Basel). Apr 08, 2024;16(7):1443. [FREE Full text] [doi: 10.3390/cancers16071443] [Medline: 38611119]

165. Patel RH, Foltz EA, Witkowski A, Ludzik J. Analysis of artificial intelligence-based approaches applied to non-invasive imaging for early detection of melanoma: a systematic review. Cancers (Basel). Sep 23, 2023;15(19):4694. [FREE Full text] [doi: 10.3390/cancers15194694] [Medline: 37835388]

166. Jairath N, Pahalyants V, Shah R, Weed J, Carucci JA, Criscito MC. Artificial intelligence in dermatology: a systematic review of its applications in melanoma and keratinocyte carcinoma diagnosis. Dermatol Surg. Sep 01, 2024;50(9):791-798. [doi: 10.1097/DSS.0000000000004223] [Medline: 38722750]

167. Thomsen K, Iversen L, Titlestad TL, Winther O. Systematic review of machine learning for diagnosis and prognosis in dermatology. J Dermatolog Treat. Aug 2020;31(5):496-510. [FREE Full text] [doi: 10.1080/09546634.2019.1682500] [Medline: 31625775]

168. Foltz EA, Witkowski A, Becker AL, Latour E, Lim JY, Hamilton A, et al. Artificial intelligence applied to non-invasive imaging modalities in identification of nonmelanoma skin cancer: a systematic review. Cancers (Basel). Feb 01, 2024;16(3):629. [FREE Full text] [doi: 10.3390/cancers16030629] [Medline: 38339380]

169. Jones OT, Matin RN, van der Schaar M, Prathivadi Bhayankaram K, Ranmuthu CKI, Islam MS, et al. Artificial intelligence and machine learning algorithms for early detection of skin cancer in community and primary care settings: a systematic review. Lancet Digit Health. Jun 2022;4(6):e466-e476. [FREE Full text] [doi: 10.1016/S2589-7500(22)00023-1] [Medline: 35623799]

170. Dildar M, Akram S, Irfan M, Khan HU, Ramzan M, Mahmood AR, et al. Skin cancer detection: a review using deep learning techniques. Int J Environ Res Public Health. May 20, 2021;18(10):5479. [FREE Full text] [doi: 10.3390/ijerph18105479] [Medline: 34065430]

171. Kassem MA, Hosny KM, Damaševičius R, Eltoukhy MM. Machine learning and deep learning methods for skin lesion classification and diagnosis: a systematic review. Diagnostics (Basel). Jul 31, 2021;11(8):1390. [FREE Full text] [doi: 10.3390/diagnostics11081390] [Medline: 34441324]

172. Salinas MP, Sepúlveda J, Hidalgo L, Peirano D, Morel M, Uribe P, et al. A systematic review and meta-analysis of artificial intelligence versus clinicians for skin cancer diagnosis. NPJ Digit Med. May 14, 2024;7(1):125. [FREE Full text] [doi: 10.1038/s41746-024-01103-x] [Medline: 38744955]

173. Ferrante di Ruffano L, Takwoingi Y, Dinnes J, Chuchu N, Bayliss SE, Davenport C, et al. Cochrane Skin Cancer Diagnostic Test Accuracy Group. Computer-assisted diagnosis techniques (dermoscopy and spectroscopy-based) for diagnosing skin cancer in adults. Cochrane Database Syst Rev. Dec 04, 2018;12(12):CD013186. [FREE Full text] [doi: 10.1002/14651858.CD013186] [Medline: 30521691]

174. Furriel BCRS, Oliveira BD, Prôa R, Paiva JQ, Loureiro RM, Calixto WP, et al. Artificial intelligence for skin cancer detection and classification for clinical environment: a systematic review. Front Med (Lausanne). 2023;10:1305954. [FREE Full text] [doi: 10.3389/fmed.2023.1305954] [Medline: 38259845]

175. Zhong J, Hu Y, Ge X, Xing Y, Ding D, Zhang G, et al. A systematic review of radiomics in chondrosarcoma: assessment of study quality and clinical value needs handy tools. Eur Radiol. Feb 2023;33(2):1433-1444. [doi: 10.1007/s00330-022-09060-3] [Medline: 36018355]

176. Zhang H, Lei H, Pang J. Diagnostic performance of radiomics in adrenal masses: A systematic review and meta-analysis. Front Oncol. Sep 2, 2022;12:975183. [FREE Full text] [doi: 10.3389/fonc.2022.975183] [Medline: 36119492]

177. Dai X, Zhao B, Zang J, Wang X, Liu Z, Sun T, et al. Diagnostic performance of radiomics and deep learning to identify benign and malignant soft tissue tumors: a systematic review and meta-analysis. Acad Radiol. Oct 2024;31(10):3956-3967. [doi: 10.1016/j.acra.2024.03.033] [Medline: 38614826]

178. Zhu N, Meng X, Wang Z, Hu Y, Zhao T, Fan H, et al. Radiomics in diagnosis, grading, and treatment response assessment of soft tissue sarcomas: a systematic review and meta-analysis. Acad Radiol. Oct 2024;31(10):3982-3992. [doi: 10.1016/j.acra.2024.03.029] [Medline: 38772802]

179. Ong W, Zhu L, Tan YL, Teo EC, Tan JH, Kumar N, et al. Application of machine learning for differentiating bone malignancy on imaging: a systematic review. Cancers (Basel). Mar 18, 2023;15(6):1837. [FREE Full text] [doi: 10.3390/cancers15061837] [Medline: 36980722]

180. Salehi MA, Mohammadi S, Harandi H, Zakavi SS, Jahanshahi A, Shahrabi Farahani M, et al. Diagnostic performance of artificial intelligence in detection of primary malignant bone tumors: a meta-analysis. J Imaging Inform Med. Apr 2024;37(2):766-777. [FREE Full text] [doi: 10.1007/s10278-023-00945-3] [Medline: 38343243]

181. Bai A, Si M, Xue P, Qu Y, Jiang Y. Artificial intelligence performance in detecting lymphoma from medical imaging: a systematic review and meta-analysis. BMC Med Inform Decis Mak. Jan 08, 2024;24(1):13. [FREE Full text] [doi: 10.1186/s12911-023-02397-9] [Medline: 38191361]

182. Bezzi C, Mapelli P, Presotto L, Neri I, Scifo P, Savi A, et al. Radiomics in pancreatic neuroendocrine tumors: methodological issues and clinical significance. Eur J Nucl Med Mol Imaging. Nov 2021;48(12):4002-4015. [doi: 10.1007/s00259-021-05338-8] [Medline: 33835220]

183. Li JO, Liu H, Ting DSJ, Jeon S, Chan RVP, Kim JE, et al. Digital technology, tele-medicine and artificial intelligence in ophthalmology: A global perspective. Prog Retin Eye Res. May 2021;82:100900. [FREE Full text] [doi: 10.1016/j.preteyeres.2020.100900] [Medline: 32898686]

184. Bisschops R, East JE, Hassan C, Hazewinkel Y, Kamiński MF, Neumann H, et al. Advanced imaging for detection and differentiation of colorectal neoplasia: European Society of Gastrointestinal Endoscopy (ESGE) Guideline - Update 2019. Endoscopy. Dec 2019;51(12):1155-1179. [FREE Full text] [doi: 10.1055/a-1031-7657] [Medline: 31711241]

185. Ntoutsi E, Fafalios P, Gadiraju U, Iosifidis V, Nejdl W, Vidal ME, et al. Bias in Data-driven AI Systems—An Introductory Survey. arXiv. 2020. URL: https://arxiv.org/abs/2001.09762 [accessed 2025-03-05]

186. Bateman RM, Sharpe MD, Jagger JE, Ellis CG, Solé-Violán J, López-Rodríguez M, et al. 36th International Symposium on Intensive Care and Emergency Medicine : Brussels, Belgium. 15-18 March 2016. Crit Care. Apr 20, 2016;20:94. [doi: 10.1186/s13054-016-1208-6] [Medline: 27885969]

187. Obermeyer Z, Emanuel EJ. Predicting the future - big data, machine learning, and clinical medicine. N Engl J Med. Sep 29, 2016;375(13):1216-1219. [FREE Full text] [doi: 10.1056/NEJMp1606181] [Medline: 27682033]

188. Nensa F, Demircioglu A, Rischpler C. Artificial intelligence in nuclear medicine. J Nucl Med. Sep 2019;60(Suppl 2):29S-37S. [FREE Full text] [doi: 10.2967/jnumed.118.220590] [Medline: 31481587]

189. Duarte-Rojo A, Sejdic E. Artificial intelligence and the risk for intuition decline in clinical medicine. Am J Gastroenterol. Mar 01, 2022;117(3):401-402. [doi: 10.14309/ajg.0000000000001618] [Medline: 35029157]

190. Yang G, Ye Q, Xia J. Unbox the black-box for the medical explainable AI via multi-modal and multi-centre data fusion: A mini-review, two showcases and beyond. Inf Fusion. Jan 2022;77:29-52. [FREE Full text] [doi: 10.1016/j.inffus.2021.07.016] [Medline: 34980946]

191. Ugga L, Perillo T, Cuocolo R, Stanzione A, Romeo V, Green R, et al. Meningioma MRI radiomics and machine learning: systematic review, quality score assessment, and meta-analysis. Neuroradiology. Aug 2021;63(8):1293-1304. [FREE Full text] [doi: 10.1007/s00234-021-02668-0] [Medline: 33649882]

192. Alabi RO, Elmusrati M, Leivo I, Almangush A, Mäkitie A. Artificial intelligence-driven radiomics in head and neck cancer: current status and future prospects. Int J Med Inform. Aug 2024;188:105464. [FREE Full text] [doi: 10.1016/j.ijmedinf.2024.105464] [Medline: 38728812]

193. Wang T, Hong J, Huang J, Liao C, Lu C, Wu Y. Systematic review and meta-analysis of deep learning applications in computed tomography lung cancer segmentation. Radiother Oncol. Aug 2024;197:110344. [doi: 10.1016/j.radonc.2024.110344] [Medline: 38806113]

194. Koohi-Moghadam M, Bae KT. Generative AI in medical imaging: applications, challenges, and ethics. J Med Syst. Aug 31, 2023;47(1):94. [doi: 10.1007/s10916-023-01987-4] [Medline: 37651022]

195. Li J, Chen J, Tang Y, Wang C, Landman BA, Zhou SK. Transforming medical imaging with Transformers? A comparative review of key properties, current progresses, and future perspectives. Med Image Anal. Apr 2023;85:102762. [FREE Full text] [doi: 10.1016/j.media.2023.102762] [Medline: 36738650]

196. Zhou H, Yu Y, Wang C, Zhang S, Gao Y, Pan J, et al. A transformer-based representation-learning model with unified processing of multimodal input for clinical diagnostics. Nat Biomed Eng. Jun 2023;7(6):743-755. [doi: 10.1038/s41551-023-01045-x] [Medline: 37308585]

197. Arora A, Arora A. The promise of large language models in health care. Lancet. Feb 25, 2023;401(10377):641. [doi: 10.1016/S0140-6736(23)00216-7] [Medline: 36841609]

198. Shah NH, Entwistle D, Pfeffer MA. Creation and adoption of large language models in medicine. JAMA. Sep 05, 2023;330(9):866-869. [doi: 10.1001/jama.2023.14217] [Medline: 37548965]

199. Thirunavukarasu AJ, Ting DSJ, Elangovan K, Gutierrez L, Tan TF, Ting DSW. Large language models in medicine. Nat Med. Aug 2023;29(8):1930-1940. [doi: 10.1038/s41591-023-02448-8] [Medline: 37460753]

200. Lotter W, Hassett MJ, Schultz N, Kehl KL, Van Allen EM, Cerami E. Artificial intelligence in oncology: current landscape, challenges, and future directions. Cancer Discov. May 01, 2024;14(5):711-726. [doi: 10.1158/2159-8290.CD-23-1199] [Medline: 38597966]

201. Malliori A, Pallikarakis N. Breast cancer detection using machine learning in digital mammography and breast tomosynthesis: A systematic review. Health Technol. Aug 12, 2022;12(5):893-910. [doi: 10.1007/s12553-022-00693-4]

202. Grossman D, Sweeney C, Doherty JA. The rapid rise in cutaneous melanoma diagnoses. N Engl J Med. Apr 08, 2021;384(14):e54. [doi: 10.1056/NEJMc2101980] [Medline: 33826834]

203. Geras KJ, Mann RM, Moy L. Artificial intelligence for mammography and digital breast tomosynthesis: current concepts and future perspectives. Radiology. Nov 2019;293(2):246-259. [FREE Full text] [doi: 10.1148/radiol.2019182627] [Medline: 31549948]

204. Amann J, Blasimme A, Vayena E, Frey D, Madai VI, Precise4Q consortium. Explainability for artificial intelligence in healthcare: a multidisciplinary perspective. BMC Med Inform Decis Mak. Nov 30, 2020;20(1):310. [FREE Full text] [doi: 10.1186/s12911-020-01332-6] [Medline: 33256715]

205. Julia H. Juggling more than three balls at once: multilevel jurisdictional challenges in EU Data Protection Regulation. International Journal of Law and Information Technology. 2019;27(2):142-170. [FREE Full text] [doi: 10.1093/ijlit/eaz002]

206. Lambin P, Leijenaar RTH, Deist TM, Peerlings J, de Jong EEC, van Timmeren J, et al. Radiomics: the bridge between medical imaging and personalized medicine. Nat Rev Clin Oncol. Dec 2017;14(12):749-762. [doi: 10.1038/nrclinonc.2017.141] [Medline: 28975929]

207. Mongan J, Moy L, Kahn CE. Checklist for artificial intelligence in medical imaging (claim): a guide for authors and reviewers. Radiol Artif Intell. Mar 2020;2(2):e200029. [FREE Full text] [doi: 10.1148/ryai.2020200029] [Medline: 33937821]

208. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. BMJ. Jan 07, 2015;350:g7594. [FREE Full text] [doi: 10.1136/bmj.g7594] [Medline: 25569120]

209. Zwanenburg A, Vallières M, Abdalah MA, Aerts HJWL, Andrearczyk V, Apte A, et al. The image biomarker standardization initiative: standardized quantitative radiomics for high-throughput image-based phenotyping. Radiology. May 2020;295(2):328-338. [FREE Full text] [doi: 10.1148/radiol.2020191145] [Medline: 32154773]

210. Whiting PF, Rutjes AWS, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2 Group. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med. Oct 18, 2011;155(8):529-536. [FREE Full text] [doi: 10.7326/0003-4819-155-8-201110180-00009] [Medline: 22007046]

211. Dercle L, McGale J, Sun S, Marabelle A, Yeh R, Deutsch E, et al. Artificial intelligence and radiomics: fundamentals, applications, and challenges in immunotherapy. J Immunother Cancer. Sep 2022;10(9):e005292. [FREE Full text] [doi: 10.1136/jitc-2022-005292] [Medline: 36180071]

212. Papadimitriou N, Markozannes G, Kanellopoulou A, Critselis E, Alhardan S, Karafousia V, et al. An umbrella review of the evidence associating diet and cancer risk at 11 anatomical sites. Nat Commun. Jul 28, 2021;12(1):4579. [FREE Full text] [doi: 10.1038/s41467-021-24861-8] [Medline: 34321471]

Abbreviations

AI: artificial intelligence
AUC: area under the curve
CLAIM: Checklist for Artificial Intelligence in Medical Imaging
CNN: convolutional neural network
CNS: central nervous system
CT: computed tomography
DL: deep learning
GRADE: Grading of Recommendations, Assessment, Development, and Evaluation
IBSI: Image Biomarker Standardization Initiative
JBI: Joanna Briggs Institute
MRI: magnetic resonance imaging
PET: positron emission tomography
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
QUADAS-2: Quality Assessment of Diagnostic Accuracy Studies-2
RQS: radiomics quality score
SVM: support vector machine
TRIPOD: Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis
US: ultrasound