International Journal of Biomedicine 16(4) (2026) 442-454
http://dx.doi.org/10.21103/Article16(4)_RA3

REVIEW ARTICLE

INTERNATIONAL JOURNAL OF BIOMEDICINE

The Clinical Application of AI in Advanced Medical Imaging and Image Processing in Diagnostic Radiology

Rowa Aljondi[1*]

[1]*Department of Applied Radiologic Technology, College of Applied Medical Sciences, University of Jeddah, Jeddah, Saudi Arabia*

Abstract

Background: Artificial intelligence (AI), particularly deep learning (DL), now interacts with every part of the imaging chain, from raw data acquisition and image reconstruction to segmentation, diagnosis, prognosis, and follow-up. Evidence has grown rapidly, but it is spread across modalities, tasks, and reporting standards, making it difficult to assess for clinical translation.

Objective: To provide an integrated, clinically focused overview of AI in advanced medical imaging and image processing, with particular emphasis on the imaging modalities magnetic resonance imaging (MRI), computed tomography (CT), and positron emission tomography (PET).

Methods and Results: Structured narrative review using the PRISMA-ScR principles. We searched PubMed/MEDLINE, Embase, Web of Science, Scopus, and IEEE Xplore for papers from January 2015 to June 2026. The two anchor reviews were Wang et al. (2021) on DL-based reconstruction and Obuchowicz et al. (2024) on clinical image-processing applications.

Deep learning reconstruction is the most advanced clinical application. Commercial and regulator-cleared products reduce MRI acquisition time by approximately 60–80%, reduce the CT dose without compromising low-contrast detectability, and reduce PET acquisition or tracer dose by 50%-75%. Self-configuring segmentation networks (such as nnU-Net) achieve Dice coefficients ≥0.90 for many downstream organs; prospective randomized trials in breast screening show that AI-dependent reading can maintain or increase cancer detection rates while reducing radiologists' workload by approximately 44%. A multimodal approach extends imaging by incorporating radiomics to predict treatment response and enable long-term monitoring. Challenges remain: lack of external validation, data shift, algorithmic bias, hallucination in generative reconstruction, poor interpretability, and fragmented regulation.

Conclusion: AI is moving from proof of concept to clinical infrastructure. Safe scale-up requires physics-informed, uncertainty-aware models; multicentre, federated validation; adherence to CLAIM, TRIPOD+AI, and DECIDE-AI; and post-deployment monitoring. **(International Journal of Biomedicine. 2026;16(4):442-454.)**

Keywords: deep learning, magnetic resonance imaging, computed tomography, positron emission tomography, clinical translation

For citation: Aljondi R. The Clinical Application of AI in Advanced Medical Imaging and Image Processing in Diagnostic Radiology. International Journal of Biomedicine. 2026;16(4):442-454. doi:10.21103/Article16(4)_RA3

Introduction

In modern medicine, medical imaging is the primary non-invasive means of accessing anatomy, physiology, and molecular function and underpins most diagnostic and therapeutic decisions. MRI provides unmatched soft-tissue contrast without ionizing radiation, CT provides fast, high-resolution attenuation maps, and PET quantifies metabolic and receptor-level processes. However, each modality has a structural trade-off: MRI has slow acquisition times, CT and PET involve radiation exposure, and the data generated by each modality is too large for clinicians to review thoroughly. However, a family of representation learning methods known as deep learning (DL), based on multilayer neural networks, has revolutionized how these trade-offs are managed. Convolutional architectures, and more recently transformer and diffusion models,[1-3] have penetrated all layers of the imaging pipeline since 2015. Early surveys identified hundreds of proof-of-concept studies in classification, detection, and segmentation,[2] and later reviews framed AI as a pathway to "high-performance medicine," where algorithms support, but do not supplant, clinical practice.[4,5] By 2023, the field had advanced to the point that a key question was whether regulators would approve and reimburse a product.[6]

© 2026 The Author(s). *International Journal of Biomedicine* is published by IMRDC.
This is an open access article under the CC BY-NC-ND 4.0 license.

There has been a parallel evolution of two strands of literature. The first is image formation: how DL reconstructs diagnostic images from undersampled MRI k-space, Low-dose CT sinograms, or Low-count PET data. Wang et al. provided a thorough overview of this strand for MRI, CT, and PET, covering model-based and data-driven perspectives, as well as early commercial products with FDA 510(k) clearance.[7] The second strand focuses on image analysis: segmentation, detection, classification, prediction, and quality control. Obuchowicz et al. compiled 34 recent studies from a variety of modalities including X-ray, MRI, CT, SPECT, ultrasonography, mammography, and histopathology, and divided clinical applications into seven categories: segmentation, detection and diagnosis, pre-processing, personalized treatment planning, predictive analytics, quality control, and monitoring and follow-up.[8]

The two strands cannot be separated in the clinic. Noise texture changes due to reconstruction decisions, which in turn affect segmentation accuracy and radiomic feature values, ultimately influencing prognosis prediction.[9–11] When evaluating a DL reconstruction product, a radiology department should first consider its impact on each downstream task, not just perceived image quality.

This review bridges the two strands. It makes specific contributions to:

1. Summarize the fundamental DL architectures and learning paradigms to be mapped by the clinical reader to imaging physics.

2. Summarize the evidence of DL reconstruction in MRI, CT, and PET, including commercial products and reported performance.

3. Map the clinical applications to the seven categories of Obuchowicz et al., including some representative quantitative results and the best available evidence (randomized trials).

4. Summarize modality-specific evidence and the regulatory landscape.

5. Discuss the technical, clinical, ethical and regulatory challenges to adoption and propose a research agenda.

Review Methodology

This is a structured narrative review. It follows the PRISMA 2020 reporting guidelines[12] and, where applicable, the PRISMA-ScR extension for scoping reviews,[13] as appropriate for a non-systematic synthesis. It was not registered and does not pool effect sizes.

Anchor sources. We used two peer-reviewed reviews to set the conceptual boundaries: Wang et al.[7] for the DL-based reconstruction of medical images, MRI/CT/PET, and Obuchowicz et al.[8] for clinical applications of image processing. They asked us to trace the primary studies back to their origins and follow articles backward.

Databases and period. We searched PubMed/MEDLINE, Embase, Web of Science Core Collection, Scopus, and IEEE Xplore for English-language articles published from 1 January 2015 to 30 June 2026. We used pre-2015 landmark papers in the methodological field (such as compressed sensing, ML-EM, and OSEM reconstruction) as background.

Search concepts. Three blocks were combined with AND: (i) *"artificial intelligence" OR "deep learning" OR "machine learning" OR "neural network" OR "transformer" OR "diffusion model" OR "foundation model"*; (ii) *"magnetic resonance imaging" OR "computed tomography" OR "positron emission tomography" OR "mammography" OR "radiography" OR "ultrasonography"*; (iii) *"reconstruction" OR "segmentation" OR "detection" OR "diagnosis" OR "radiomics" OR "prognosis" OR "treatment planning" OR "quality control"*.

Selection. We prioritized randomized controlled trials and prospective studies over other study designs, followed by multicentre or externally validated retrospective studies, systematic reviews and meta-analyses, seminal method papers that established an architecture or paradigm, and reporting guidelines and regulatory analyses and included only single-centre studies without external validation in the absence of external validation when they depicted a distinct clinical task. We excluded editorials, conference abstracts without full text, and non-imaging AI.

Fundamentals of AI for Medical Imaging

From handcrafted features to learned representations

There are two families of AI-based imaging methods.[8] Classical machine learning (ML) relies on a priori features, including intensity statistics, shape descriptors, and texture. Radiomics formalizes this by viewing images as 'minable data' rather than just pictures.[14,15] Later, the Image Biomarker Standardization Initiative (IBSI) standardized 169 feature definitions, improving feature reproducibility across software and centers.[16,17]

Deep learning, in contrast, extracts hierarchical features directly from pixels or raw signals: low-level layers find edges and textures, whereas higher levels find semantic structure.[1,3] Deep learning scores are now higher on most tasks, but not on every task. In fact, Nurzynska et al. showed that, despite using a small dataset, a handcrafted feature based on texture analysis of vertebral bodies on CT predicted patient age better than DL, [18] suggesting that small datasets and well-known physics can still favor handcrafted features.

Learning paradigms

Three paradigms are relevant.[7] *Supervised learning* minimizes the loss between predictions and paired ground truth (i.e., low-dose and full-dose CT pairs). When full-dose references or fully sampled data are unavailable, *unsupervised and self-supervised learning* leverage data structure. *Reinforcement learning* provides a reward signal to train an agent; it is seldom used, for example, to tune iterative reconstruction parameters automatically. Most reconstruction/analysis models used in the clinic are supervised, with self-supervised pre-training emerging as a possibility.

Core architectures

Table 1 summarizes the architectures that recur throughout this review. There are two trends here. First, self-configuring frameworks like nnU-Net show that an effective pipeline can matter more than innovative architecture, as nnU-Net matched or surpassed most specialized approaches

Table 1.

Core deep learning architectures in medical imaging and their main clinical roles.

Architecture	Key idea	Main imaging roles	References
Convolutional neural network (CNN)	Local, weight-shared filters learn spatial features hierarchically	Classification, detection, denoising	[1,2]
U-Net	Encoder–decoder with skip connections preserves spatial detail	Segmentation; image-to-image reconstruction	[19]
Residual network (ResNet)	Identity shortcuts let very deep networks train stably	Backbone for classification; residual denoising	[20]
Generative adversarial network (GAN)	Generator and discriminator compete, producing realistic images	Low-dose denoising, super-resolution, cross-modality synthesis	[21]
nnU-Net	Self-configuring U-Net pipeline adapts preprocessing, architecture and training to each dataset	Out-of-the-box segmentation across organs and modalities	[22]
Vision transformer / UNETR	Self-attention models long-range context	3D segmentation, multimodal fusion	[23,24]
Diffusion model	Learns to reverse a gradual noising process	Reconstruction, synthesis, data augmentation	[25]
Foundation model	Large model pre-trained on broad data, adapted to many tasks	Promptable segmentation (SAM, MedSAM); generalist diagnosis	[26–30]

on 23 public datasets.[22] Second, foundation models trained on billions of images guarantee label-efficient adaptation. For example, MedSAM, trained on over 1.5 million image–mask pairs, segments structures from a simple prompt,2[7]and a self-supervised retinal model, RETFound, generalizes to predicting both ocular and systemic disease.[29] A new trend is the development of generalist medical AI models that integrate images, text, and other information.[28,30]

Training and evaluation essentials

Good models require a training, validation, and test set, preferably collected from a different institution or scanner.[7] To avoid overfitting, data is enhanced using techniques such as data augmentation (affine transforms, noise injection). Typically, reconstruction models are trained using mean squared error (L2) or mean absolute error (L1), or a combination of structural similarity (SSIM), perceptual, or adversarial loss.[7] Performance measurement should not be limited to pixel measures such as peak signal-to-noise ratio (PSNR) and SSIM but should also include task-based measures such as lesion detectability, segmentation agreement, and reader studies.

AI in Image Reconstruction: MRI, CT, and PET

Reconstruction is the first step in the clinical workflow, and its quality carries over to all subsequent steps in the computational pipeline.[7] Acquisition time limits MRI; radiation dose or counting statistics limit CT and PET. DL methods are tackling these challenges in three ways: (i) image domain post-processing – a network refines the initial analytic reconstruction; (ii) model-based (unrolled) networks – iterations of a physics-based optimization are unfolded into network layers with data-consistency steps; and (iii) direct (end-to-end) mapping from raw data to image.

Magnetic resonance imaging

Classical MRI must obey the Nyquist sampling in k-space, slowing dynamic and high-resolution protocols. Parallel imaging and compressed sensing (CS) relax this requirement; CS relies on image sparsity in a transform domain[31] but requires handcrafted priors and long iterative solvers. In MR reconstruction, Wang et al. first applied DL to map undersampled data to fully sampled images in 2016.[32] AUTOMAP then showed that a network could learn the domain transform directly from raw sensor data to image.[33]

Model-based networks soon became the dominant design because they contained the MR forward model. These works all involved unrolling a gradient-descent scheme as a variational network: Hammernik et al.[34] for multi-coil knee MRI;Schlemper et al.[35] for a deep cascade of CNN blocks combined with data-consistency layers for dynamic cardiac MRI; MoDL used a learned denoiser and data-consistency step based on the conjugate gradient algorithm for dynamic cardiac MRI.[36] Data-driven GANs enhanced perceptual sharpness in CS-MRI.[37]

The key was publicly available raw data. The fastMRI dataset was released with raw k-space and DICOM images of the knee and, later, the brain,[38] and the 2020 fastMRI challenge showed that the best models produced high-radiological-quality images at 4-fold acceleration, whereas subtle pathology was lost at 8-fold acceleration,[39] For fully sampled references, self-supervised physics-guided training eliminates the need for a reference,[40] Comprehensive surveys of parallel imaging summarize these trends,[41] and distribution-based diffusion models are increasingly used as powerful imaging-acceleration priors in MRI.[42]

A growing body of clinical evidence supports this. A prospective interchangeability study found 4-fold DL-accelerated knee MRI at 3 T to be diagnostically interchangeable with conventional imaging for diagnosing internal derangement.[43] A commercial DL product achieved 60% faster volumetric brain MRI without compromising automated volumetric measurements.[44] Failure monitoring and radiologist training are important translation issues that must be addressed, along with protocol harmonization.[45]

Computed tomography

The CT reconstruction has progressed from Filtered back-projection (FBP) to Statistical iterative reconstruction and model-based iterative reconstruction and now to DL reconstruction (DL reconstruction).[46,47] Iterative methods suppress noise at low dose, but may create a "plastic" noise texture and are often computationally expensive.

This field was sparked by the 2016 AAPM Low-Dose CT Grand Challenge.[48] Some initial networks included the residual encoder–decoder CNN (RED-CNN),[49] a directional-wavelet CNN,5[1] and FBPConvNet (residual U-Net),[51] which corrected FBP images. The sparse-view CT was solved using the framelet-based U-Net[52] and the DD-Net.[53] Adversarial and perceptual losses preserved the texture better than pixel-wise losses for GANs.[54,55] A modularized deep net (MDN) achieved results comparable to or better than those of commercial iterative reconstruction (IR) in a blinded reader study.[56] Theories are reviewed, and the trend toward physics-integrated designs [57,58] is summarized.

DLR has been commercialized, and clinical studies support this. DLR has been shown to enhance the image quality of ultra-high-resolution abdominal CT compared with hybrid and model-based iterative reconstruction,[59] and phantom studies have measured the potential for dose reduction using a DL algorithm.[60] In head CT, the SNR improved by up to 54% in grey matter and 60% in white matter, and the CNR improved by up to 58% compared with adaptive statistical iterative reconstruction (ASIR-V).[61] Lower noise levels and higher CNR in unenhanced brain CT images were achieved using DL-based reconstruction compared with hybrid iterative reconstruction.[62] Research algorithms have continued to expand the range: the use of shape priors for GAN-based sinogram inpainting boosted PSNR for limited-angle CT by 7 dB,[63] and an approximation to iterative reconstruction (UnetU) for photon-counting micro-CT enabled a much faster material decomposition.[64]

Positron emission tomography

Analytic FBP, iterative maximum-likelihood expectation-maximization (MLEM),[65] and accelerated iterative ordered-subsets EM (OSEM)[66] have been used for PET reconstruction. The clinical drivers for AI are shorter acquisition times (to minimize motion and increase throughput) and a lower tracer dose (to minimize radiation, particularly in children and with repeated exams).

In 2017, Xiang et al.[67] combined DL with PET to estimate standard-dose images from low-dose PET and to co-register MRI and PET using a deep auto-context CNN. Gong et al. incorporated a CNN representation into iterative reconstruction,[68] while DeepPET reconstructed images directly from sinograms using an encoder–decoder network, a much faster approach than OSEM.[69] A 3D conditional GAN generated full-dose PET images from one-quarter-dose PET.[70] The DL with multi-contrast MRI inputs yielded diagnostic-quality images from ultra-low-dose (around 1%) 18F-florbetaben PET for amyloid imaging,[71] and a GAN with task-specific perceptual loss was able to do the same without MRI.[72]These methods are discussed in detail by Reader et al.[73]

In clinical applications, this enabled a 2-fold reduction in 18F-FDG acquisition time without sacrificing diagnostic performance when using the digital PET/CT platform for DL denoising,[74] and a positive business case was generated for real-world implementation of AI-based reduced-dose FDG PET.[75]

Cross-modality synthesis and attenuation correction

PET/MRI does not have a direct measure of photon attenuation. For brain PET/MRI, pseudo-CT generated by DL showed better performance than the atlas and Dixon methods in achieving correct attenuation correction,[76] and zero-echo-time and Dixon inputs also showed good performance for correct attenuation correction in the pelvis.[77] This synthesis principle applies to MR-only radiotherapy planning.[78]

Commercial products and reported performance

Not every vendor-reported gain is replicated independently, and there are differences between jurisdictions in what indications are considered "cleared" (Table 2).

Dose and scan time should be reduced only after confirming performance locally with phantom and patient studies.

Stability, hallucination and trust

The reconstructed DL is unstable: very small changes in the input produce large artifacts, and adding more training

Table 2.

Representative regulator-cleared DL reconstruction products and reported benefits, as summarised by Wang et al. [7] and subsequent clinical studies.

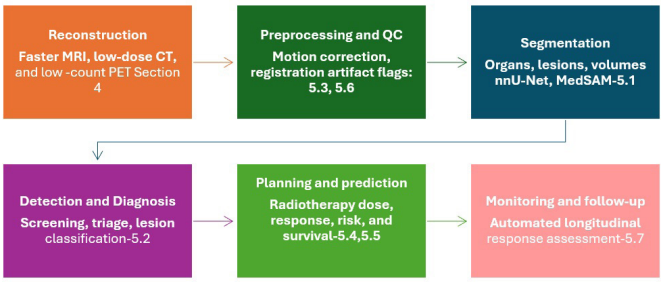
Product (vendor)	Modality	Reported benefit	References
AI-assisted compressed sensing, ACS (United Imaging)	MRI	About 80% scan-time reduction for most fast spin-echo sequences; most body-part protocols under 100 s	[7]
SubtleMR (Subtle Medical)	MRI	About 60% scan-time reduction; volumetric measures preserved	[7,44]
AiCE (Canon Medical)	CT	Lower noise, higher signal and contrast-to-noise ratio than hybrid iterative reconstruction	[7,59,62]
TrueFidelity (GE HealthCare)	CT	Improved image quality and dose-reduction potential at low dose	[7,60,61]
DELTA (United Imaging)	CT	Up to 80% dose reduction; low-contrast detectability improved by up to 157%	[7]
HYPER DLR (United Imaging)	PET	Acquisition shortened from 3 to under 1 min per bed position; SNR improved by 42%	[7]
SubtlePET (Subtle Medical)	PET	Diagnostic images from 25–50% of the standard acquisition time	[7,74,75]

samples does not necessarily improve the result.[79] Hallucination maps and uncertainty estimates motivate generative models to hallucinate realistic but false structures. These risks can be mitigated with data consistency layers, physics-based unrolling, and task-based evaluation (lesion detectability and not PSNR only).

Clinical Applications Across the Imaging Pathway

Obuchowicz et al. categorized clinical AI apps into seven categories.[8] The same structure is used, with each category linked to the best available evidence. In the following, these categories are located along the patient's imaging pathway in Figure 1.

Figure 1. The seven application categories of Obuchowicz et al.[8] and the reconstruction framework of Wang et al.[7] are incorporated into the clinical imaging pathway to formulate AI applications. Later steps have varying levels of evidence, with the most clinically mature step being reconstruction.



Highlighted: the most clinically mature step, with regulator – cleared products in MRI, CT, and PET

Fig. 1. *AI applications along the imaging pathway · 6 stages.*

Image segmentation

Segmentation, which is the delineation of organs, lesions, or tissue classes, is the basis for the following: volumetry, radiotherapy planning, radiomics, and surgical navigation. Considerable progress has been made in recent years through public challenges such as the Brain Tumor Segmentation (BraTS) challenge[81] and the Medical Segmentation Decathlon (MSD), which includes 10 CT and MRI tasks.[82] nnU-Net is the winner of the Decathlon and is still a good baseline.[22,82] TotalSegmentator achieved a high Dice coefficient of 0.943 on routine CT, and a fully convolutional network for cardiac MRI was evaluated with human-level accuracy in nearly 5,000 participants from the UK Biobank.[84]

In oncology, nnU-Net was found to outperform other architectures for prostate gland and zonal segmentation on MRI[85] and to segment pediatric Wilms tumor with a median Dice of 0.90 and a median Hausdorff distance of 7.2 mm compared to manual segmentation.[86] Model-based segmentation can be extended to new structures with only limited annotations, called promptable foundation models.[27] The choice of metric is important: Boundary errors or missed small lesions can be covered up by dice alone, and the Metrics Reloaded framework supplies task-specific advice.[87] In radiotherapy, the segmentation of 21 organs at risk was at the expert level for the head and neck.[88]

Disease detection and diagnosis

Detection and diagnosis are the most researched applications. A meta-analysis of 24 studies showed a mean sensitivity of 89% for brain metastases on an MRI scan; however, there was a lack of homogeneity with respect to false-positive reports, which prevented a consolidated analysis.[89]

Breast imaging is evidence-based to the highest degree. A DL system decreased the number of false positives (FP) by 1.2% and 5.7% and the number of false negatives (FN) by 2.7% and 9.4%, respectively, in retrospective UK and US data.[90] Stand-alone AI was comparable to the average of 101 radiologists[91] and, in an external evaluation, the top three commercially available algorithms showed improved sensitivity over the first-reader radiologists.[92] Prospective evidence followed. In the MASAI trial (over 80,000 women), AI-assisted screening was found not to affect screening workload or false-positive rates adversely and led to a 44% reduction in cancer detection.[93] The nationwide implementation study PRAIM in Germany (463,094 women) with AI-supported double reading confirmed a higher detection rate compared to standard double reading, with no increase in the recall rate.[95] In the ScreenTrustCAD study, AI plus 1 radiologist was proven to be non-inferior to standard double reading.[94] DL also has the potential to enhance the interpretation of digital breast tomosynthesis (DBT):[96] a single-slice DBT classifier achieved an accuracy of 93.2%,[97] and an MRI model, when applied to BRCA carriers, was able to retrospectively identify 65% of cancers earlier than radiologists.[98]

Thoracic imaging: an end-to-end model for low-dose CT lung screening achieved an AUC of 0.944 and, when applied to nontraining scans, reduced false positives by 11% and false negatives by 5% compared with six radiologists.[99] CheXNeXt identified 14 pathologies on chest radiographs, achieving a 100% match rate with the radiologists.[100] A YOLO-based detector on chest radiographs has achieved an average precision of 0.815[101] in detecting foreign objects.

Focused abdominal ultrasound is used for emergency and abdominal imaging, and automated ML is used to classify haemoperitoneum, achieving sensitivity and specificity greater than 94% and an AUROC of 0.97 on external validation.[102] AI tools that facilitate large-vessel occlusion diagnosis on CT angiography reduce stroke workflow,[103] and multinational CT models have achieved approximately 90% accuracy for COVID-19 pneumonia.[104] CNN has classified liver masses on multiphase CT,[105] and a multiparametric MRI pipeline has distinguished healthy livers from livers with lesions, with an AUC of 0.85.[106] An MRI of the prostate can be used to identify clinically significant cancer with a multiple-instance support vector machine (SVM) with a 92% accuracy rate.[107]

In addition to radiology, the diabetic retinopathy study,[108] dermatology,[109] and lung histopathology[110] demonstrated the performance of specialists and provided methodological models that the radiology community has adopted.

Image preprocessing and enhancement

Preprocessing enhances the data fed into human readers and subsequent algorithms. In addition to the reconstruction methods presented in Section 4, applications include unsupervised

deformable registration (VoxelMorph) that registers volumes in under a minute, versus under an hour for classical methods,[112] and retrospective correction of the signal loss caused by motion in diffusion-weighted MRI of the liver, which improved ADC accuracy in the left lobe.[111] Scan times are also shortened by using synthetic contrast generation and super-resolution, provided the outputs are validated against acquired images.

Personalized treatment planning

Automatic contouring is now routinely used in many radiotherapy centers, and has been shown to limit planning time and inter-observer variability.[113] For prostate intensity-modulated radiotherapy[114] and glioblastoma volumetric modulated arc therapy, DL models have been applied directly to the anatomy to predict 3D dose distributions, showing sensitivity to anatomical variations and even supporting quality assurance of automatic contouring[115] in glioblastoma. Image-based DL has also been applied to individualize radiotherapy dose for lung cancer,[116] and MR-only planning is also supported by MR-derived synthetic CT.[78]

Predictive analytics

Predictive models integrate imaging with clinical data to predict prognosis, treatment response/likelihood of future complications. A DL network was able to predict survival in non-small-cell lung cancer (NSCLC) from routine computed tomography (CT) across multiple cohorts.[117] Another study of locally advanced cervical cancer showed that a DL model using multiparametric MRI outperformed clinical models when predicting response to neoadjuvant chemotherapy.[118] DL risk models based on mammography were superior to the Tyrer–Cuzick model,[119] and Sybil predicted future lung cancer from a single low-dose CT with an AUC of up to 0.92 over 1 year,[120] and chest radiographs alone predicted long-term mortality.[121] In general, fusion of imaging with electronic health records and molecular data significantly outperforms the single-modality models.[122-124]

Quality control

Before the image is presented to the reader or an algorithm, the automated quality control checks the image first. A DL model agreed with radiologists in its assessment of the diagnostic quality of liver MRI on T2-weighted images, [125] and reference-free networks have been developed to detect motion artifacts in MRI, thereby allowing repeated sequences. At the same time, the patient remains on the table.[126] Quality control also involves identifying out-of-distribution input on which an AI model is likely to fail.

Monitoring and follow-up

Automated and repeatable measurement is a benefit of longitudinal assessment. An artificial neural network was used in a multicentre neuro-oncology study to quantify tumor response on MRI and to predict overall survival more accurately than the central neuro-oncology radiological assessment, based on the Response Assessment in Neuro-Oncology (RANO) criteria[127] (Table 3). Pathological complete response in rectal cancer was predicted using multi-task DL on serial MRI.[128] The same applies to pediatric follow-up, e.g., through automated volumetry using nnU-Net-based Wilms tumor measurement.[86]

Modality-Specific Evidence and the Regulatory Landscape

Maturity by modality

The evidence base is uneven across modalities. Table 4 summarises where AI has reached routine practice and where it remains investigational.

Regulatory clearance and evidence gaps

Regulatory approval has clinical trial results. Many of the AI devices that make their way to the European market do so before the US market, with the largest proportion of cleared devices in radiology.[129] Retrospective evaluation was used by nearly all of the 130 FDA-cleared AI devices, and

Table 3.

Representative clinical applications and quantitative results.

Category	Task (modality)	Method	Key result	References
Segmentation	104 structures (CT)	nnU-Net-based	Mean Dice 0.943	[83]
Segmentation	Wilms tumor (MRI)	nnU-Net	Median Dice 0.90; HD95 7.2 mm	[86]
Detection	Brain metastases (MRI)	DL meta-analysis	Pooled sensitivity 89%	[89]
Detection	Breast screening (mammography)	AI-supported reading, RCT	Detection maintained or increased; workload −44%	[93]
Detection	Breast screening (mammography)	AI-supported double reading	Detection +17.6%; recall not increased	[95]
Detection	Lung screening (LDCT)	3D CNN	AUC 0.944	[99]
Diagnosis	Hemoperitoneum (US)	AutoML	AUROC 0.97 (external)	[102]
Diagnosis	Prostate cancer (MRI)	Texture + MIL-SVM	Accuracy 92%	[107]
Preprocessing	Low-dose CT	GAN inpainting	PSNR +7 dB	[63]
Treatment planning	Glioblastoma dose (CT)	Cascaded 3D U-Net	Contour-sensitive dose prediction	[115]
Prediction	Future lung cancer (LDCT)	Sybil	1-year AUC up to 0.92	[120]
Monitoring	Glioma response (MRI)	ANN	Better survival prediction than central RANO assessment	[127]

Table 4.
Maturity of AI applications by modality.

Modality	Most mature applications	Highest level of evidence	Main open problems	References
MRI	DL reconstruction for accelerated scans; brain, prostate and cardiac segmentation	Prospective interchangeability studies; regulator-cleared products	Hallucination at high acceleration; vendor- and field-strength dependence	[33–45,84,85]
CT	DL reconstruction for low dose; whole-body segmentation; lung nodule and stroke triage	Reader studies; multicenter retrospective validation	Noise-texture shift affects radiomics; generalization across scanners	[48–64,83,99]
PET and PET/MRI	Low-count denoising; fast acquisition; MR-based attenuation correction	Clinical non-inferiority studies; real-world deployments	Quantitative bias in standardized uptake values; tracer diversity	[67–77]
Mammography and DBT	Screening support; triage; workload reduction	Randomized and nationwide prospective studies	Interval-cancer outcomes; equity across populations	[90–98]
Radiography	Chest radiograph triage; foreign-body and fracture detection	Multireader retrospective studies	Label noise; shortcut learning	[100,101,121]
Ultrasound	Focused emergency examinations; automated measurements	Multicenter retrospective validation	Operator dependence; video-level analysis	[102]
SPECT and nuclear medicine	Dementia staging; denoising	Single-center studies	Small datasets	[8]

most were not assessed across multiple sites.[130] Sixty-four of 100 commercially available CE-marked radiology products had no peer-reviewed evidence of efficacy, and only a minority had evidence of clinical impact beyond diagnostic accuracy.[131]

Now the regulations are becoming stricter. To guide post-clearance adaptive AI devices, the FDA has established predetermined change-control plans in the U.S. On top of the Medical Device Regulation, obligations for data governance, transparency, and human oversight have been introduced, and most medical AI is classified as high-risk under the European Union's Artificial Intelligence Act. The Saudi Food and Drug Authority (SFDA) has published specific guidance for the review and approval of AI- and ML-based medical devices, aligning with the Saudi Vision 2030 digital health objectives. The reality for departments is that it is important to start, rather than finish, the evaluation process—in addition to clearance, prospective monitoring and periodic revalidation are required.

Challenges, Limitations and Future Directions

Evidence quality and generalizability

One of the main flaws in the field is the disconnect between retrospective accuracy and real-world benefit.[132] Of the 516 published imaging AI studies, only 6% performed external validation.[133] When external validation does occur, it is usually accompanied by a drop in algorithm performance, with approximately a quarter of studies showing a significant decline.[134] There are shortcuts that models can learn: pneumonia detectors trained on chest radiographs learned to detect the hospital of origin rather than the disease.[135]

Data: quantity, quality, shift and privacy

There is limited, curated, well-labeled, and representative data.[136] Domain shift arises from variations in scanners, protocols, reconstruction kernels, and populations. It can be alleviated by domain adaptation,[137] synthetic data generated by generative models,[138] and harmonization, but these methods are not a substitute for training data with diversity. Federated learning is a way to train models across institutions without moving patient data. It achieved 99% of the performance of a centralized model for brain tumor segmentation,[139] and a 20-institution federated model for predicting COVID-19 outcomes showed significant improvement in generalisability.[140] The scalability of this approach relies on privacy-preserving computation, secure aggregation, and governance mechanisms.[141,142] In the Gulf region, especially in Saudi Arabia, federated and regionally curated datasets are particularly significant, as most public training data come from North American, European, and East Asian populations.

Bias and equity

AI can embed and exacerbate health inequities. Systematic underdiagnosis of disease was observed using state-of-the-art chest radiograph classifiers in female, Black, Hispanic, and younger patients, as well as in patients with Medicaid insurance.[143] DL models have also been shown to predict a patient's race from medical images, even for degraded images, which experts cannot.[144] Subgroup performance reporting and fairness audits should therefore be standard.

Interpretability and trust

Many models of reconstruction and analysis of DLs are black boxes.[7] The instability of saliency methods like Grad-CAM, which produce heat maps to pinpoint parts of an image that influence a prediction,[145] and their inability to explain their importance are well known. Radiology-specific reviews emphasize the need to appraise explanations with clinical users.[146,147] The physics-based unrolled networks, along with the data-consistency layers and pixel-wise uncertainty maps, provide more meaningful transparency into reconstruction than post hoc saliency.[79,80]

Reporting standards and evaluation frameworks

There are guidelines for the AI life cycle now: CLAIM for imaging studies in AI [148] (updated 2024[149]); TRIPOD+AI for prediction models;[150] CONSORT-AI for randomized trials;[151] and an international consensus on trustworthy and deployable medical AI (FUTURE-AI).[153] Journals and funders should

insist on compliance, and departments should consider these structures when purchasing commercial tools. The need is obvious: systematic reviews have shown that DL can achieve accuracy comparable to clinicians on internal test sets, while prospective studies are limited and many are at high risk of bias.[154–156] Authors, reviewers, and readers have practical appraisal guides to use.[157]

Workflow integration, regulation and liability

Regardless of the tools' accuracy, they can fail if they generate clicks or alerts, or cause delays. For successful deployment, these elements must be integrated with PACS and RIS; there must be clear guidelines on who is responsible for taking action on AI output; and there must be ongoing monitoring for any performance drift.[45] Issues of liability (manufacturer, institution or radiologist) and reimbursement are not fully clear.[6]

Future directions

Five directions appear most promising:

1. Physics-informed and self-supervised reconstruction that maintains data consistency, quantifies uncertainty, and requires no fully sampled or full-dose references.[40,42,58] Foundational and generalist models that can be adapted to new tasks, modalities, and institutions with few labels, and that combine images with reports and clinical data.[26-30,123]

3. Integrated pipelines from acquisition to decision that optimize the reconstruction process for subsequent tasks, such as detection or quantification, not just visual quality.[7,60]

4. Studies that have followed the MASAI and PRAIM models[93,95] in measuring patient outcomes, cost-effectiveness, and equity.

5. Fairness audits and post-market surveillance for federated, regionally representative data

Conclusions

AI, an integral part of the clinical infrastructure, now enhances medical imaging. It is most developed for reconstruction – MRI acquisitions are shortened, CT dose is reduced, and PET is accelerated, while maintaining diagnostic information, using DL methods approved by regulators. Self-configuring segmentation systems, validated in randomized breast-screening trials; predictive models based on imaging and clinical data; and automated quality control and monitoring now cover the entire imaging pathway outlined by Wang et al.[7] and Obuchowicz et al.[8]

The primary constraints have become ones other than algorithmic accuracy on curated test sets. They are generalizations across scanners and populations, not subject to hallucination or instability, fair, interpretable, and limited to prospective outcome data. To close these gaps, there is a need for physics-informed and uncertainty-aware models; for multicenter and federated validation, including under-represented populations (such as Saudi Arabia and the Gulf region); for consistent use of CLAIM, TRIPOD+AI, CONSORT-AI, and DECIDE-AI; and for ongoing post-deployment monitoring. When these conditions are met, AI can provide faster, safer, more accurate and more equitable imaging for patients.

Conflicts of Interest

The author declares no conflicts of interest.

References

1. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature 2015;52:436–444. doi: 10.1038/nature14539
2. Litjens G, Kooi T, Ehteshami Bejnordi B, Setio AAA, Ciompi F, Ghafoorian M, et al. A survey on deep learning in medical image analysis. Med Image Anal 2017;42:60–88. doi: 10.1016/j.media.2017.07.005
3. Shen D, Wu G, Suk HI. Deep learning in medical image analysis. Annu Rev Biomed Eng 2017;19:221–248. doi:10.1146/annurev-bioeng-071516-044442
4. Hosny A, Parmar C, Quackenbush J, Schwartz LH, Aerts HJWL. Artificial intelligence in radiology. Nat Rev Cancer 2018;18(8):500–510. doi: 10.1038/s41568-018-0016-5
5. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat Med 2019;25:44–56. doi: 10.1038/s41591-018-0300-7
6. Rajpurkar P, Lungren MP. The current and future state of AI interpretation of medical images. N Engl J Med 2023;388(21):1981–1990. doi: 10.1056/NEJMra2301725
7. Wang S, Cao G, Wang Y, Liao S, Wang Q, Shi J, Li C, Shen D. Review and prospect: artificial intelligence in advanced medical imaging. Front Radiol 2021;1:781868. doi: 10.3389/fradi.2021.781868
8. Obuchowicz R, Strzelecki M, Piórkowski A. Clinical applications of artificial intelligence in medical imaging and image processing—a review. Cancers 2024;16(10):1870. doi: 10.3390/cancers16101870
9. Esteva A, Chou K, Yeung S, Naik N, Madani A, Mottaghi A, et al. Deep learning-enabled medical computer vision. NPJ Digit Med 2021;4:5. doi: 10.1038/s41746-020-00376-2
10. 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 2019;69(2):127–157. doi:10.3322/caac.21552
11. Sahiner B, Pezeshk A, Hadjiiski LM, Wang X, Drukker K, Cha KH, et al. Deep learning in medical imaging and radiation therapy. Med Phys 2019;46(1):e1–e36. doi: 10.1002/mp.13264
12. 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 2021;372:n71. doi: 10.1136/bmj.n71
13. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018;169(7):467–473. doi:10.7326/M18-0850
14. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. Radiology 2016;278(2):563–577. doi: 10.1148/radiol.2015151169
15. 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 2017;14:749–762. doi: 10.1038/nrclinonc.2017.141
16. 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 2020;295(2):328–338. https://doi.org/10.1148/radiol.2020191145

17. Mayerhoefer ME, Materka A, Langs G, Häggström I, Szczypiński P, Gibbs P, et al. Introduction to radiomics. J Nucl Med 2020;61(4):488–495. doi: 10.2967/jnumed.118.222893
18. Nurzynska K, Piórkowski A, Strzelecki M, Kociołek M, Banyś RP, Obuchowicz R. Differentiating age and sex in vertebral body CT scans – texture analysis versus deep learning approach. Biocybern Biomed Eng 2024;44(1):20–30. doi: 10.1016/j.bbe.2023.11.002
19. Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation. In: Medical Image Computing and Computer-Assisted Intervention - MICCAI 2015. Lecture Notes in Computer Science, vol. 9351. Cham: Springer; 2015. p. 234–241. doi: 10.1007/978-3-319-24574-4_28
20. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 2016. p. 770–778. doi: 10.1109/CVPR.2016.90
21. Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, et al. Generative adversarial networks. Commun ACM 2020;63(11):139–144. doi: 10.1145/3422622
22. Isensee F, Jaeger PF, Kohl SAA, Petersen J, Maier-Hein KH. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. Nat Methods 2021;18:203–211. doi: 10.1038/s41592-020-01008-z
23. Hatamizadeh A, Tang Y, Nath V, Yang D, Myronenko A, Landman B, et al. UNETR: transformers for 3D medical image segmentation. In: Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV); 2022. p. 1748–1758. doi: 10.1109/WACV51458.2022.00181
24. Shamshad F, Khan S, Zamir SW, Khan MH, Hayat M, Khan FS, et al. Transformers in medical imaging: a survey. Med Image Anal 2023;88:102802. doi: 10.1016/j.media.2023.102802
25. Kazerouni A, Aghdam EK, Heidari M, Azad R, Fayyaz M, Hacihaliloglu I, et al. Diffusion models in medical imaging: a comprehensive survey. Med Image Anal 2023;88:102846. doi: 10.1016/j.media.2023.102846
26. Kirillov A, Mintun E, Ravi N, Mao H, Rolland C, Gustafson L, et al. Segment anything. In: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV); 2023. p. 3992–4003. doi: 10.1109/ICCV51070.2023.00371
27. Ma J, He Y, Li F, Han L, You C, Wang B. Segment anything in medical images. Nat Commun 2024;15:654. doi: 10.1038/s41467-024-44824-z
28. Moor M, Banerjee O, Abad ZSH, Krumholz HM, Leskovec J, Topol EJ, Rajpurkar P. Foundation models for generalist medical artificial intelligence. Nature 2023;616:259–265. doi: 10.1038/s41586-023-05881-4
29. Zhou Y, Chia MA, Wagner SK, Ayhan MS, Williamson DJ, Struyven RR, et al. A foundation model for generalizable disease detection from retinal images. Nature 2023;622:156–163. doi: 10.1038/s41586-023-06555-x
30. Tu T, Azizi S, Driess D, Schaekermann M, Amin M, Chang PC, et al. Towards generalist biomedical AI. NEJM AI 2024;1(3):AIoa2300138. https://doi.org/10.1056/AIoa2300138
31. Villota S, Inga E. Sparse Transform and Compressed Sensing Methods to Improve Efficiency and Quality in Magnetic Resonance Medical Imaging. Sensors 2025. 25(16):5137. doi: 10.3390/s25165137
32. Wang S, Su Z, Ying L, Peng X, Zhu S, Liang F, et al. Accelerating magnetic resonance imaging via deep learning. In: Proceedings of the IEEE 13th International Symposium on Biomedical Imaging (ISBI); 2016:514–517. doi: 10.1109/ISBI.2016.7493320
33. Zhu B, Liu JZ, Cauley SF, Rosen BR, Rosen MS. Image reconstruction by domain-transform manifold learning. Nature 2018;555:487–492. doi: 10.1038/nature25988
34. Hammernik K, Klatzer T, Kobler E, Recht MP, Sodickson DK, Pock T, et al. Learning a variational network for reconstruction of accelerated MRI data. Magn Reson Med 2018;79(6):3055–3071. doi: 10.1002/mrm.26977
35. Schlemper J, Caballero J, Hajnal JV, Price AN, Rueckert D. A deep cascade of convolutional neural networks for dynamic MR image reconstruction. IEEE Trans Med Imaging 2018;37(2):491–503. doi: 10.1109/TMI.2017.2760978
36. Aggarwal HK, Mani MP, Jacob M. MoDL: model-based deep learning architecture for inverse problems. IEEE Trans Med Imaging 2019;38(2):394–405. doi: 10.1109/TMI.2018.2865356
37. Yang G, Yu S, Dong H, Slabaugh G, Dragotti PL, Ye X, et al. DAGAN: deep de-aliasing generative adversarial networks for fast compressed sensing MRI reconstruction. IEEE Trans Med Imaging 2018;37(6):1310–1321. doi: 10.1109/TMI.2017.2785879
38. Knoll F, Zbontar J, Sriram A, Muckley MJ, Bruno M, Defazio A, et al. fastMRI: A publicly available raw k-space and DICOM dataset of knee images for accelerated MR image reconstruction using machine learning. Radiol Artif Intell 2020;2(1):e190007. doi: 10.1148/ryai.2020190007
39. Muckley MJ, Riemenschneider B, Radmanesh A, Kim S, Jeong G, Ko J, et al. Results of the 2020 fastMRI challenge for machine learning MR image reconstruction. IEEE Trans Med Imaging 2021;40(9):2306–2317. doi: 10.1109/TMI.2021.3075856
40. Yaman B, Hosseini SAH, Moeller S, Ellermann J, Uğurbil K, Akçakaya M. Self-supervised learning of physics-guided reconstruction neural networks without fully sampled reference data. Magn Reson Med 2020;84(6):3172–3191. doi: 10.1002/mrm.28378
41. Knoll F, Hammernik K, Zhang C, Moeller S, Pock T, Sodickson DK, et al. Deep-learning methods for parallel magnetic resonance imaging reconstruction: a survey of the current approaches, trends, and issues. IEEE Signal Process Mag 2020;37(1):128–140. doi: 10.1109/MSP.2019.2950640
42. Chung H, Ye JC. Score-based diffusion models for accelerated MRI. Med Image Anal 2022;80:102479. doi: 10.1016/j.media.2022.102479
43. Recht MP, Zbontar J, Sodickson DK, Knoll F, Yakubova N, Sriram A, et al. Using deep learning to accelerate knee MRI at 3 T: results of an interchangeability study. AJR Am J Roentgenol 2020;215(6):1421–1429. doi: 10.2214/AJR.20.23313
44. Bash S, Wang L, Airriess C, Zaharchuk G, Gong E, Shankaranarayanan A, et al. Deep learning enables 60% accelerated volumetric brain MRI while preserving quantitative performance: a prospective, multicenter, multireader trial. AJNR Am J Neuroradiol 2021;42(12):2130–2137. doi: 10.3174/ajnr.A7358
45. Chaudhari AS, Sandino CM, Cole EK, Larson DB, Gold GE, Vasanawala SS, et al. Prospective deployment of deep learning in MRI: a framework for important considerations, challenges, and recommendations for best practices. J Magn Reson Imaging 2021;54(2):357–371. doi: 10.1002/jmri.27331
46. Willemink MJ, Noël PB. The evolution of image

reconstruction for CT-from filtered back projection to artificial intelligence. Eur Radiol 2019;29(5):2185–2195. doi: 10.1007/s00330-018-5810-7
47. Wang G. A perspective on deep imaging. IEEE Access 2016;4:8914–8924. doi: 10.1109/ACCESS.2016.2624938
48. McCollough CH, Bartley AC, Carter RE, Chen B, Drees TA, Edwards P, et al. Low-dose CT for the detection and classification of metastatic liver lesions: results of the 2016 Low Dose CT Grand Challenge. Med Phys 2017;44(10):e339–e352. doi: 10.1002/mp.12345
49. Chen H, Zhang Y, Kalra MK, Lin F, Chen Y, Liao P, et al. Low-dose CT with a residual encoder-decoder convolutional neural network. IEEE Trans Med Imaging 2017;36(12):2524–2535. doi: 10.1109/TMI.2017.2715284
50. Kang E, Min J, Ye JC. A deep convolutional neural network using directional wavelets for low-dose X-ray CT reconstruction. Med Phys 2017;44(10):e360–e375. doi: 10.1002/mp.12344
51. Jin KH, McCann MT, Froustey E, Unser M. Deep convolutional neural network for inverse problems in imaging. IEEE Trans Image Process 2017;26(9):4509–4522. doi: 10.1109/TIP.2017.2713099
52. Han Y, Ye JC. Framing U-Net via deep convolutional framelets: application to sparse-view CT. IEEE Trans Med Imaging 2018;37(6):1418–1429. doi: 10.1109/TMI.2018.2823768
53. Zhang Z, Liang X, Dong X, Xie Y, Cao G. A sparse-view CT reconstruction method based on combination of DenseNet and deconvolution. IEEE Trans Med Imaging 2018;37(6):1407–1417. doi: 10.1109/TMI.2018.2823338
54. Wolterink JM, Leiner T, Viergever MA, Išgum I. Generative adversarial networks for noise reduction in low-dose CT. IEEE Trans Med Imaging 2017;36(12):2536–2545. doi: 10.1109/TMI.2017.2708987
55. Yang Q, Yan P, Zhang Y, Yu H, Shi Y, Mou X, et al. Low-dose CT image denoising using a generative adversarial network with Wasserstein distance and perceptual loss. IEEE Trans Med Imaging 2018;37(6):1348–1357. doi: 10.1109/TMI.2018.2827462
56. Shan H, Padole A, Homayounieh F, Kruger U, Khera RD, Nitiwarangkul C, et al. Competitive performance of a modularized deep neural network compared to commercial algorithms for low-dose CT image reconstruction. Nat Mach Intell 2019;1:269–276. doi: 10.1038/s42256-019-0057-9
57. Koetzier LR, Mastrodicasa D, Szczykutowicz TP, van der Werf NR, Wang AS, Sandfort V, et al. Deep learning image reconstruction for CT: technical principles and clinical prospects. Radiology 2023;306(3):e221257. doi: 10.1148/radiol.221257
58. Wang G, Ye JC, De Man B. Deep learning for tomographic image reconstruction. Nat Mach Intell 2020;2(12):737–748. doi: 10.1038/s42256-020-00273-z
59. Akagi M, Nakamura Y, Higaki T, Narita K, Honda Y, Zhou J, et al. Deep learning reconstruction improves image quality of abdominal ultra-high-resolution CT. Eur Radiol 2019;29(11):6163–6171. doi: 10.1007/s00330-019-06170-3
60. Greffier J, Hamard A, Pereira F, Barrau C, Pasquier H, Beregi JP, et al. Image quality and dose reduction opportunity of deep learning image reconstruction algorithm for CT: a phantom study. Eur Radiol 2020;30:3951–3959. doi: 10.1007/s00330-020-06724-w
61. Pula M, Kucharczyk E, Zdanowicz A, Guzinski M. Image quality improvement in deep learning image reconstruction of head computed tomography examination. Tomography 2023;9(4):1485–1493. doi: 10.3390/tomography9040118
62. Cozzi A, Cè M, De Padova G, Libri D, Caldarelli N, Zucconi F, et al. Deep learning-based versus iterative image reconstruction for unenhanced brain CT: a quantitative comparison of image quality. Tomography 2023;9(5):1629–1637. doi: 10.3390/tomography9050130
63. Valat E, Farrahi K, Blumensath T. Sinogram inpainting with generative adversarial networks and shape priors. Tomography 2023;9(3):1137–1152. doi: 10.3390/tomography9030094
64. Nadkarni R, Clark DP, Allphin AJ, Badea CT. A deep learning approach for rapid and generalizable denoising of photon-counting micro-CT images. Tomography 2023;9(4):1286–1302. doi: 10.3390/tomography9040102
65. Kasai R, Otsuka, H. Graph-Enhanced Expectation Maximization for Emission Tomography. Journal of Imaging 2026;12(1):48. doi: 10.3390/jimaging12010048
66. Huang H-M, Hsiao I-T (2016). Accelerating an Ordered-Subset Low-Dose X-Ray Cone Beam Computed Tomography Image Reconstruction with a Power Factor and Total Variation Minimization. PLoS ONE 11(4): e0153421. doi: 10.1371/journal.pone.0153421
67. Xiang L, Qiao Y, Nie D, An L, Lin W, Wang Q, Shen D. Deep auto-context convolutional neural networks for standard-dose PET image estimation from low-dose PET/MRI. Neurocomputing 2017;267:406–416. doi: 10.1016/j.neucom.2017.06.048
68. Gong K, Guan J, Kim K, Zhang X, Yang J, Seo Y, et al. Iterative PET image reconstruction using convolutional neural network representation. IEEE Trans Med Imaging 2019;38(3):675–685. doi: 10.1109/TMI.2018.2869871
69. Häggström I, Schmidtlein CR, Campanella G, Fuchs TJ. DeepPET: a deep encoder–decoder network for directly solving the PET image reconstruction inverse problem. Med Image Anal 2019;54:253–262. doi: 10.1016/j.media.2019.03.013
70. Wang Y, Yu B, Wang L, Zu C, Lalush DS, Lin W, et al. 3D conditional generative adversarial networks for high-quality PET image estimation at low dose. NeuroImage 2018;174:550–562. doi: 10.1016/j.neuroimage.2018.03.045
71. Chen KT, Gong E, de Carvalho Macruz FB, Xu J, Boumis A, Khalighi M, et al. Ultra–low-dose ^{18}F-florbetaben amyloid PET imaging using deep learning with multi-contrast MRI inputs. Radiology 2019;290(3):649–656. doi: 10.1148/radiol.2018180940
72. Ouyang J, Chen KT, Gong E, Pauly J, Zaharchuk G. Ultra-low-dose PET reconstruction using generative adversarial network with feature matching and task-specific perceptual loss. Med Phys 2019;46(8):3555–3564. doi: 10.1002/mp.13626
73. Reader AJ, Corda G, Mehranian A, da Costa-Luis C, Ellis S, Schnabel JA. Deep learning for PET image reconstruction. IEEE Trans Radiat Plasma Med Sci 2021;5(1):1–25. doi: 10.1109/TRPMS.2020.3014786
74. Weyts K, Lasnon C, Ciappuccini R, Lequesne J, Corroyer-Dulmont A, Quak E, et al. Artificial intelligence-based PET denoising could allow a two-fold reduction in [^{18}F]FDG PET acquisition time in digital PET/CT. Eur J Nucl Med Mol Imaging 2022;49(11):3750–3760. doi: 10.1007/s00259-022-05800-1
75. Katsari K, Penna D, Arena V, Polverari G, Ianniello A, Italiano D, et al. Artificial intelligence for reduced dose 18F-FDG PET examinations: a real-world deployment through a standardized framework and business case assessment.

EJNMMI Phys 2021;8:25. doi: 10.1186/s40658-021-00374-7
76. Liu F, Jang H, Kijowski R, Bradshaw T, McMillan AB. Deep learning MR imaging–based attenuation correction for PET/MR imaging. Radiology 2018;286(2):676–684. doi: 10.1148/radiol.2017170700
77. Leynes AP, Yang J, Wiesinger F, Kaushik SS, Shanbhag DD, Seo Y, et al. Zero-echo-time and Dixon deep pseudo-CT (ZeDD CT): direct generation of pseudo-CT images for pelvic PET/MRI attenuation correction using deep convolutional neural networks with multiparametric MRI. J Nucl Med 2018;59(5):852–858. doi: 10.2967/jnumed.117.198051
78. Han X. MR-based synthetic CT generation using a deep convolutional neural network method. Med Phys 2017;44(4):1408–1419. doi: 10.1002/mp.12155
79. Antun V, Renna F, Poon C, Adcock B, Hansen AC. On instabilities of deep learning in image reconstruction and the potential costs of AI. Proc Natl Acad Sci U.S.A 2020;117(48):30088–30095. doi: 10.1073/pnas.1907377117
80. Bhadra S, Kelkar VA, Brooks FJ, Anastasio MA. On hallucinations in tomographic image reconstruction. IEEE Trans Med Imaging 2021;40(11):3249–3260. doi: 10.1109/TMI.2021.3077857
81. Menze BH, Jakab A, Bauer S, Kalpathy-Cramer J, Farahani K, Kirby J, et al. The multimodal brain tumor image segmentation benchmark (BRATS). IEEE Trans Med Imaging 2015;34(10):1993–2024. doi: 10.1109/TMI.2014.2377694
82. Antonelli M, Reinke A, Bakas S, Farahani K, Kopp-Schneider A, Landman BA, et al. The Medical Segmentation Decathlon. Nat Commun 2022;13:4128. doi: 10.1038/s41467-022-30695-9
83. Wasserthal J, Breit HC, Meyer MT, Pradella M, Hinck D, Sauter AW, et al. TotalSegmentator: robust segmentation of 104 anatomic structures in CT images. Radiol Artif Intell 2023;5(5):e230024. doi: 10.1148/ryai.230024
84. Bai W, Sinclair M, Tarroni G, Oktay O, Rajchl M, Vaillant G, et al. Automated cardiovascular magnetic resonance image analysis with fully convolutional networks. J Cardiovasc Magn Reson 2018;20(1):65. doi: 10.1186/s12968-018-0471-x
85. Rodrigues NM, Silva S, Vanneschi L, Papanikolaou N. A comparative study of automated deep learning segmentation models for prostate MRI. Cancers 2023;15(5):1467. doi: 10.3390/cancers15051467
86. Buser MAD, van der Steeg AFW, Wijnen MHWA, Fitski M, van Tinteren H, van den Heuvel-Eibrink MM, et al. Radiologic versus segmentation measurements to quantify Wilms tumor volume on MRI in pediatric patients. Cancers 2023;15(7):2115. doi: 10.3390/cancers15072115
87. Maier-Hein L, Reinke A, Godau P, Tizabi MD, Buettner F, Christodoulou E, et al. Metrics reloaded: recommendations for image analysis validation. Nat Methods 2024;21:195–212. doi: 10.1038/s41592-023-02151-z
88. Nikolov S, Blackwell S, Zverovitch A, Mendes R, Livne M, De Fauw J, et al. Clinically applicable segmentation of head and neck anatomy for radiotherapy: deep learning algorithm development and validation study. J Med Internet Res 2021;23(7):e26151. doi: 10.2196/26151
89. Ozkara BB, Chen MM, Federau C, Karabacak M, Briere TM, Li J, Wintermark M. Deep learning for detecting brain metastases on MRI: a systematic review and meta-analysis. Cancers 2023;15(2):334. doi: 10.3390/cancers15020334
90. McKinney SM, Sieniek M, Godbole V, Godwin J, Antropova N, Ashrafian H, et al. International evaluation of an AI system for breast cancer screening. Nature 2020;577:89–94. doi: 10.1038/s41586-019-1799-6
91. 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 2019;111(9):916–922. doi: 10.1093/jnci/djy222
92. Salim M, Wåhlin E, Dembrower K, Azavedo E, Foukakis T, Liu Y, et al. External evaluation of 3 commercial artificial intelligence algorithms for independent assessment of screening mammograms. JAMA Oncol 2020;6(10):1581–1588. doi: 10.1001/jamaoncol.2020.3321
93. Lång K, Josefsson V, Larsson AM, Larsson S, Högberg C, Sartor H, et al. Artificial intelligence-supported screen reading versus standard double reading in the Mammography Screening with Artificial Intelligence trial (MASAI): a clinical safety analysis of a randomised, controlled, non-inferiority, single-blinded, screening accuracy study. Lancet Oncol 2023;24(8):936–944. doi: 10.1016/S1470-2045(23)00298-X
94. Dembrower K, Crippa A, Colón E, Eklund M, Strand F; ScreenTrustCAD Trial Consortium. Artificial intelligence for breast cancer detection in screening mammography in Sweden: a prospective, population-based, paired-reader, non-inferiority study. Lancet Digit Health 2023;5(10):e703–e711. doi: 10.1016/S2589-7500(23)00153-X
95. Eisemann N, Bunk S, Mukama T, Baltus H, Elsner SA, Gomille T, et al. Nationwide real-world implementation of AI for cancer detection in population-based mammography screening. Nat Med 2025;31:917–924. doi: 10.1038/s41591-024-03408-6
96. Lotter W, Diab AR, Haslam B, Kim JG, Grisot G, Wu E, et al. Robust breast cancer detection in mammography and digital breast tomosynthesis using an annotation-efficient deep learning approach. Nat Med 2021;27:244–249. doi: 10.1038/s41591-020-01174-9
97. Mendes J, Matela N, Garcia N. Avoiding tissue overlap in 2D images: single-slice DBT classification using convolutional neural networks. Tomography 2023;9(1):398–412. doi: 10.3390/tomography9010032
98. Anaby D, Shavin D, Zimmerman-Moreno G, Nissan N, Friedman E, Sklair-Levy M. 'Earlier than early' detection of breast cancer in Israeli BRCA mutation carriers applying AI-based analysis to consecutive MRI scans. Cancers 2023;15(12):3120. doi: 10.3390/cancers15123120
99. Ardila D, Kiraly AP, Bharadwaj S, Choi B, Reicher JJ, Peng L, et al. End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. Nat Med 2019;25:954–961. doi: 10.1038/s41591-019-0447-x
100. Rajpurkar P, Irvin J, Ball RL, Zhu K, Yang B, Mehta H, et al. Deep learning for chest radiograph diagnosis: a retrospective comparison of the CheXNeXt algorithm to practicing radiologists. PLoS Med 2018;15(11):e1002686. doi: 10.1371/journal.pmed.1002686
101. Kufel J, Bargieł-Łączek K, Koźlik M, Czogalik Ł, Dudek P, Magiera M, et al. Chest X-ray foreign objects detection using artificial intelligence. J Clin Med 2023;12(18):5841. doi: 10.3390/jcm12185841
102. Jeong D, Jeong W, Lee JH, Park SY. Use of automated machine learning for classifying hemoperitoneum on ultrasonographic images of Morrison's pouch: a multicenter retrospective study. J Clin Med 2023;12(12):4043. doi: 10.3390/jcm12124043
103. Murray NM, Unberath M, Hager GD, Hui FK.

Artificial intelligence to diagnose ischemic stroke and identify large vessel occlusions: a systematic review. J Neurointerv Surg 2020;12(2):156–164. doi:10.1136/neurintsurg-2019-015135
104. Harmon SA, Sanford TH, Xu S, Turkbey EB, Roth H, Xu Z, et al. Artificial intelligence for the detection of COVID-19 pneumonia on chest CT using multinational datasets. Nat Commun 2020;11:4080. doi: 10.1038/s41467-020-17971-2
105. Yasaka K, Akai H, Abe O, Kiryu S. Deep learning with convolutional neural network for differentiation of liver masses at dynamic contrast-enhanced CT: a preliminary study. Radiology 2018;286(3):887–896. doi: 10.1148/radiol.2017170706
106. Skwirczyński M, Tabor Z, Lasek J, Schneider Z, Gibała S, Kucybała I, et al. Deep learning algorithm for differentiating patients with a healthy liver from patients with liver lesions based on MR images. Cancers 2023;15(12):3142. doi: 10.3390/cancers15123142
107. Gibala S, Obuchowicz R, Lasek J, Schneider Z, Piórkowski A, Pociask E, Nurzynska K. Textural features of MR images correlate with an increased risk of clinically significant cancer in patients with high PSA levels. J Clin Med 2023;12(8):2836. doi: 10.3390/jcm12082836
108. Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA 2016;316(22):2402–2410. doi: 10.1001/jama.2016.17216
109. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, Thrun S. Dermatologist-level classification of skin cancer with deep neural networks. Nature 2017;542:115–118. doi: 10.1038/nature21056
110.Coudray N, Ocampo PS, Sakellaropoulos T, Narula N, Snuderl M, Fenyö D, et al. Classification and mutation prediction from non–small cell lung cancer histopathology images using deep learning. Nat Med 2018;24:1559–1567. doi: 10.1038/s41591-018-0177-5
111.Raspe J, Harder FN, Rupp S, McTavish S, Peeters JM, Weiss K, et al. Retrospective motion artifact reduction by spatial scaling of liver diffusion-weighted images. Tomography 2023;9(5):1839–1856. doi: 10.3390/tomography9050146
112.Balakrishnan G, Zhao A, Sabuncu MR, Guttag J, Dalca AV. VoxelMorph: a learning framework for deformable medical image registration. IEEE Trans Med Imaging 2019;38(8):1788–1800. doi: 10.1109/TMI.2019.2897538
113.Cardenas CE, Yang J, Anderson BM, Court LE, Brock KB. Advances in auto-segmentation. Semin Radiat Oncol 2019;29(3):185–197. doi: 10.1016/j.semradonc.2019.02.001
114.Nguyen D, Long T, Jia X, Lu W, Gu X, Iqbal Z, Jiang S. A feasibility study for predicting optimal radiation therapy dose distributions of prostate cancer patients from patient anatomy using deep learning. Sci Rep 2019;9:1076. doi: 10.1038/s41598-018-37741-x
115.Poel R, Kamath AJ, Willmann J, Andratschke N, Ermiş E, Aebersold DM, et al. Deep-learning-based dose predictor for glioblastoma–assessing the sensitivity and robustness for dose awareness in contouring. Cancers 2023;15(17):4226. doi: 10.3390/cancers15174226
116.Lou B, Doken S, Zhuang T, Wingerter D, Gidwani M, Mistry N, et al. An image-based deep learning framework for individualising radiotherapy dose: a retrospective analysis of outcome prediction. Lancet Digit Health 2019;1(3):e136–e147. doi: 10.1016/S2589-7500(19)30058-5
117.Hosny A, Parmar C, Coroller TP, Grossmann P, Zeleznik R, Kumar A, et al. Deep learning for lung cancer prognostication: a retrospective multi-cohort radiomics study. PLoS Med 2018;15(11):e1002711. doi: 10.1371/journal.pmed.1002711
118.Zhang Y, Wu C, Xiao Z, Lv F, Liu Y. A deep learning radiomics nomogram to predict response to neoadjuvant chemotherapy for locally advanced cervical cancer: a two-center study. Diagnostics 2023;13(6):1073. doi: 10.3390/diagnostics13061073
119.Yala A, Lehman C, Schuster T, Portnoi T, Barzilay R. A deep learning mammography-based model for improved breast cancer risk prediction. Radiology 2019;292(1):60–66. doi: 10.1148/radiol.2019182716
120. Mikhael PG, Wohlwend J, Yala A, Karstens L, Xiang J, Takigami AK, et al. Sybil: a validated deep learning model to predict future lung cancer risk from a single low-dose chest computed tomography. J Clin Oncol 2023;41(12):2191–2200. doi: 10.1200/JCO.22.01345
121. Lu MT, Ivanov A, Mayrhofer T, Hosny A, Aerts HJWL, Hoffmann U. Deep learning to assess long-term mortality from chest radiographs. JAMA Netw Open 2019;2(7):e197416. doi: 10.1001/jamanetworkopen.2019.7416
122. Huang SC, Pareek A, Seyyedi S, Banerjee I, Lungren MP. Fusion of medical imaging and electronic health records using deep learning: a systematic review and implementation guidelines. NPJ Digit Med 2020;3:136. doi: 10.1038/s41746-020-00341-z
123. Acosta JN, Falcone GJ, Rajpurkar P, Topol EJ. Multimodal biomedical AI. Nat Med 2022;28:1773–1784. doi: 10.1038/s41591-022-01981-2
124. Lipkova J, Chen RJ, Chen B, Lu MY, Barbieri M, Shao D, et al. Artificial intelligence for multimodal data integration in oncology. Cancer Cell 2022;40(10):1095–1110. doi: 10.1016/j.ccell.2022.09.012
125. Esses SJ, Lu X, Zhao T, Shanbhogue K, Dane B, Bruno M, Chandarana H. Automated image quality evaluation of T_2-weighted liver MRI utilizing deep learning architecture. J Magn Reson Imaging 2018;47(3):723–728. doi: 10.1002/jmri.25779
126. Küstner T, Liebgott A, Mauch L, Martirosian P, Bamberg F, Nikolaou K, et al. Automated reference-free detection of motion artifacts in magnetic resonance images. Magn Reson Mater Phy 2018;31:243–256. doi: 10.1007/s10334-017-0650-z
127. Kickingereder P, Isensee F, Tursunova I, Petersen J, Neuberger U, Bonekamp D, et al. Automated quantitative tumour response assessment of MRI in neuro-oncology with artificial neural networks: a multicentre, retrospective study. Lancet Oncol 2019;20(5):728–740. doi: 10.1016/S1470-2045(19)30098-1
128. Jin C, Yu H, Ke J, Ding P, Yi Y, Jiang X, et al. Predicting treatment response from longitudinal images using multi-task deep learning. Nat Commun 2021;12:1851. doi: 10.1038/s41467-021-22188-y
129. Muehlematter UJ, Daniore P, Vokinger KN. Approval of artificial intelligence and machine learning-based medical devices in the USA and Europe (2015–20): a comparative analysis. Lancet Digit Health 2021;3(3):e195–e203. doi: 10.1016/S2589-7500(20)30292-2
130. Wu E, Wu K, Daneshjou R, Ouyang D, Ho DE, Zou J. How medical AI devices are evaluated: limitations and recommendations from an analysis of FDA approvals. Nat Med 2021;27:582–584. doi: 10.1038/s41591-021-01312-x
131. 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 2021;31:3797–3804. doi: 10.1007/s00330-021-07892-z
132. Kelly CJ, Karthikesalingam A, Suleyman M, Corrado G, King D. Key challenges for delivering clinical impact with artificial intelligence. BMC Med 2019;17:195. doi: 10.1186/s12916-019-1426-2
133. Kim DW, Jang HY, Kim KW, Shin Y, Park SH. Design characteristics of studies reporting the performance of artificial intelligence algorithms for diagnostic analysis of medical images: results from recently published papers. Korean J Radiol 2019;20(3):405–410. doi: 10.3348/kjr.2019.0025
134. Yu AC, Mohajer B, Eng J. External validation of deep learning algorithms for radiologic diagnosis: a systematic review. Radiol Artif Intell 2022;4(3):e210064. doi: 10.1148/ryai.210064
135. Zech JR, Badgeley MA, Liu M, Costa AB, Titano JJ, Oermann EK. Variable generalization performance of a deep learning model to detect pneumonia in chest radiographs: a cross-sectional study. PLoS Med 2018;15(11):e1002683. doi: 10.1371/journal.pmed.1002683
136. Willemink MJ, Koszek WA, Hardell C, Wu J, Fleischmann D, Harvey H, et al. Preparing medical imaging data for machine learning. Radiology 2020;295(1):4–15. doi: 10.1148/radiol.2020192224
137. Guan H, Liu M. Domain adaptation for medical image analysis: a survey. IEEE Trans Biomed Eng 2022;69(3):1173–1185. doi: 10.1109/TBME.2021.3117407
138. Chen RJ, Lu MY, Chen TY, Williamson DFK, Mahmood F. Synthetic data in machine learning for medicine and healthcare. Nat Biomed Eng 2021;5:493–497. doi: 10.1038/s41551-021-00751-8
139. Sheller MJ, Edwards B, Reina GA, Martin J, Pati S, Kotrotsou A, et al. Federated learning in medicine: facilitating multi-institutional collaborations without sharing patient data. Sci Rep 2020;10:12598. doi: 10.1038/s41598-020-69250-1
140. Dayan I, Roth HR, Zhong A, Harouni A, Gentili A, Abidin AZ, et al. Federated learning for predicting clinical outcomes in patients with COVID-19. Nat Med 2021;27:1735–1743. doi: 10.1038/s41591-021-01506-3
141. Rieke N, Hancox J, Li W, Milletarì F, Roth HR, Albarqouni S, et al. The future of digital health with federated learning. NPJ Digit Med 2020;3:119. doi: 10.1038/s41746-020-00323-1
142. Kaissis GA, Makowski MR, Rückert D, Braren RF. Secure, privacy-preserving and federated machine learning in medical imaging. Nat Mach Intell 2020;2:305–311. doi: 10.1038/s42256-020-0186-1
143. Seyyed-Kalantari L, Zhang H, McDermott MBA, Chen IY, Ghassemi M. Underdiagnosis bias of artificial intelligence algorithms applied to chest radiographs in under-served patient populations. Nat Med 2021;27:2176–2182. doi: 10.1038/s41591-021-01595-0
144. Gichoya JW, Banerjee I, Bhimireddy AR, Burns JL, Celi LA, Chen LC, et al. AI recognition of patient race in medical imaging: a modelling study. Lancet Digit Health 2022;4(6):e406–e414. doi: 10.1016/S2589-7500(22)00063-2
145. Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. Grad-CAM: visual explanations from deep networks via gradient-based localization. In: Proceedings of the IEEE International Conference on Computer Vision (ICCV); 2017. p. 618–626. doi: 10.1109/ICCV.2017.74
146. Reyes M, Meier R, Pereira S, Silva CA, Dahlweid FM, von Tengg-Kobligk H, et al. On the interpretability of artificial intelligence in radiology: challenges and opportunities. Radiol Artif Intell 2020;2(3):e190043. doi: 10.1148/ryai.2020190043
147. van der Velden BHM, Kuijf HJ, Gilhuijs KGA, Viergever MA. Explainable artificial intelligence (XAI) in deep learning-based medical image analysis. Med Image Anal 2022;79:102470. doi: 10.1016/j.media.2022.102470
148. Mongan J, Moy L, Kahn CE Jr. Checklist for Artificial Intelligence in Medical Imaging (CLAIM): a guide for authors and reviewers. Radiol Artif Intell 2020;2(2):e200029. doi: 10.1148/ryai.2020200029
149. Tejani AS, Klontzas ME, Gatti AA, Mongan JT, Moy L, Park SH, et al. Checklist for Artificial Intelligence in Medical Imaging (CLAIM): 2024 update. Radiol Artif Intell 2024;6(4):e240300. doi: 10.1148/ryai.240300
150. Collins GS, Moons KGM, Dhiman P, Riley RD, Beam AL, Van Calster B, et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. BMJ 2024;385:e078378. doi: 10.1136/bmj-2023-078378
151. Liu X, Cruz Rivera S, Moher D, Calvert MJ, Denniston AK; SPIRIT-AI and CONSORT-AI Working Group. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. Nat Med 2020;26:1364–1374. doi: 10.1038/s41591-020-1034-x
152. Vasey B, Nagendran M, Campbell B, Clifton DA, Collins GS, Denaxas S, et al. Reporting guideline for the early-stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI. Nat Med 2022;28:924–933. doi: 10.1038/s41591-022-01772-9
153. Lekadir K, Frangi AF, Porras AR, Glocker B, Cintas C, Langlotz CP, et al. FUTURE-AI: international consensus guideline for trustworthy and deployable artificial intelligence in healthcare. BMJ 2025;388:e081554. doi: 10.1136/bmj-2024-081554
154. Liu X, Faes L, Kale AU, Wagner SK, Fu DJ, Bruynseels A, et al. A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis. Lancet Digit Health 2019;1(6):e271–e297. doi: 10.1016/S2589-7500(19)30123-2
155. 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 2021;4:65. doi:10.1038/s41746-021-00438-z
156. Nagendran M, Chen Y, Lovejoy CA, Gordon AC, Komorowski M, Harvey H, et al. Artificial intelligence versus clinicians: systematic review of design, reporting standards, and claims of deep learning studies. BMJ 2020;368:m689. doi: 10.1136/bmj.m689
157. Bluemke DA, Moy L, Bredella MA, Ertl-Wagner BB, Fowler KJ, Goh VJ, et al. Assessing radiology research on artificial intelligence: a brief guide for authors, reviewers, and readers-from the Radiology Editorial Board. Radiology 2020;294(3):487–489. doi: 10.1148/radiol.2019192515

***Corresponding author**: Dr. Rowa Aljondi. E-mail: rmaljondi@uj.edu.sa*