

Review

Not peer-reviewed version

---

# Clinical AI in Radiology: Foundations, Trends, and Emerging Directions with Use Cases from Moffitt Cancer Center

---

[Iryna Hartsock](#)\*, [Nikolas Koutsoubis](#), [Sabeen Ahmed](#), [Nathan Parker](#), [Matthew Schabath](#), [Cyrillo Araujo](#), [Aliya Qayyum](#), [Cesar Lam](#), [Robert Gatenby](#), [Ghulam Rasool](#)

Posted Date: 6 February 2026

doi: 10.20944/preprints202602.0555.v1

Keywords: radiology; artificial intelligence; Moffitt Cancer Center





Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Review

# Clinical AI in Radiology: Foundations, Trends, and Emerging Directions with Use Cases from Moffitt Cancer Center

Iryna Hartsock<sup>1,\*</sup> , Nikolas Koutsoubis<sup>1,2</sup>, Sabeen Ahmed<sup>1</sup>, Nathan Parker<sup>3</sup>, Matthew Schabath<sup>4</sup>, Cyrillo Araujo<sup>1,5,6</sup>, Aliya Qayyum<sup>1,5,6</sup>, Cesar Lam<sup>1,5,6</sup>, Robert Gatenby<sup>5,6</sup> and Ghulam Rasool<sup>1,5,6,7</sup> 

<sup>1</sup> Department of Machine Learning, H. Lee Moffitt Cancer Center & Research Institute

<sup>2</sup> Department of Electrical Engineering, University of South Florida

<sup>3</sup> Department of Health Outcomes and Behavior, H. Lee Moffitt Cancer Center & Research Institute

<sup>4</sup> Department of Cancer Epidemiology, H. Lee Moffitt Cancer Center & Research Institute

<sup>5</sup> Department of Diagnostic Imaging & Interventional Radiology, H. Lee Moffitt Cancer Center & Research Institute

<sup>6</sup> Department of Oncologic Sciences, University of South Florida

<sup>7</sup> Department of Neuro-Oncology, H. Lee Moffitt Cancer Center & Research Institute

\* Correspondence: iryna.hartsock@moffitt.org

**Abstract:** Artificial intelligence (AI) is at the vanguard of transforming radiology in several ways, including augmenting diagnoses, improving workflows, and increasing operational efficiency. Several integration challenges, including concerns over privacy, clinical usability, and workflow compatibility, still remain. This review discusses the foundations and current trends of clinical AI in radiology to provide essential context for ongoing developments. We then outline four key use cases from the Moffitt Cancer Center: (1) local deployment of large language models (LLMs) for restructuring and streamlining radiology reports, improving clarity and consistency without relying on external resources; (2) multimodal AI frameworks combining CT images, clinical data, laboratory biomarkers, and LLM-extracted features from clinical notes for early detection of cachexia in pancreatic cancer; (3) privacy-preserving federated learning (FL) infrastructure enabling collaborative AI model development across institutions without sharing raw patient data; and (4) an uncertainty-aware de-identification pipeline for removing Protected Health Information (PHI) from radiology images and clinical reports to support secure data analysis and sharing. We further discuss emerging opportunities for tumor board decision support, clinical trial matching, radiology report quality assurance, and the development of an imaging complexity index. Our experience highlights the critical importance of local deployment, multimodal reasoning, privacy preservation, and human-in-the-loop oversight in translating AI models from research to oncology radiology practice.

**Keywords:** radiology; artificial intelligence; Moffitt Cancer Center

## 1. Introduction

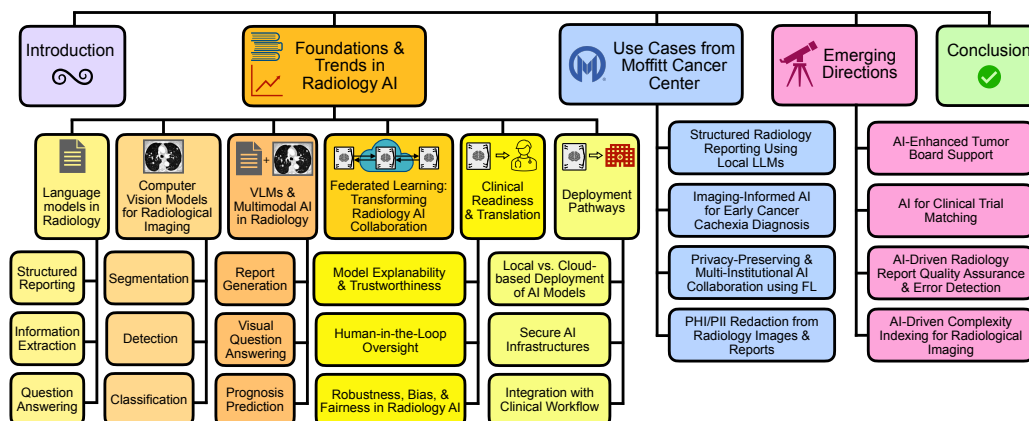
Radiology plays a central role in oncology, including early detection, diagnosis, staging, treatment monitoring, and post-therapy surveillance. As medical imaging volumes grow and case complexity increases, radiologists face mounting demands for accurate and rapid interpretation, efficient reporting, and multidisciplinary communication [1]. Artificial intelligence (AI) offers a promising solution to support radiologists across these workflows [2]. Large language models (LLMs), such as GPT-4/5, Vicuna, Rad-Phi2 have been applied to radiology-specific tasks, including structured reporting, automated impression filling, extraction of key diagnostic information from unstructured text, and question answering [3–5]. Vision-based models, like U-Net, YOLO, and ResNet, have been increasingly used in radiology image segmentation, detection, and classification tasks [6–8]. Advances in multimodal learning and vision language models (VLMs) have introduced new capabilities for

combining imaging information with clinical text to enhance diagnostic reasoning and support informed decision-making. Models such as Flamingo CXR (2024) [9], RadVLM (2025) [10], Med3D VLM (2025) [11], RadAlign (2025) [12], and RadZero (2025) [13] enable tasks including automated report generation, visual question answering (VQA), and prognosis prediction, highlighting their potential in oncology and beyond [14,15]. Alongside these systems, several radiology foundation models (FMs) have begun to emerge. RadFM (2023) [16], CURIA (2025) [17], and ONCOPILOT (2025) [18] provide broader, transferable representations that can support a wide range of downstream radiology and oncology tasks, reflecting a shift toward more general purpose radiology AI. As multimodal models continue to grow in scale, standardized preprocessing pipelines are essential for consistent image quality and reproducible model performance. Frameworks such as Pillar-0 provide modular tools for large scale preprocessing, intensity normalization, and dataset curation, which support the development of both unimodal and multimodal radiology FMs [19].

Despite rapid technical advances, many radiology teams remain cautious about integrating AI tools into daily practice [2]. AI development often occurs in research environments and may not always align with the realities of clinical radiology workflows [20]. Concerns over opaque “black-box” decision-making remain a major barrier to clinical trust [21]. As a result, models may lack the usability, reliability, generalizability, and transparency required at the point-of-care, with diagnostic accuracy issues and the risk of automation bias further limiting adoption. In oncology imaging specifically, where complex, multi-modality assessments and nuanced interpretations are routine, the absence of cancer-focused, radiologist-centered AI tools exacerbates these challenges [22]. Privacy regulations and institutional data governance frameworks may restrict the use of cloud-based solutions, while some models remain proprietary and unsuitable for customization, or local deployment. Consequently, AI adoption in radiology remains measured and fragmented, despite radiologists’ clear interest in tools that can meaningfully reduce reporting burden, enhance reader confidence, and ultimately improve patient care.

At the H. Lee Moffitt Cancer Center & Research Institute, we emphasize AI solutions that directly address radiology needs through secure, local deployment, and clinician-centered designs. Rather than adopting generic commercial platforms, our approach focuses on adapting state-of-the-art methods to the specific challenges of cancer imaging, including the high prevalence of multi-modality studies, the importance of longitudinal follow-up in oncology care, and the need for clear communication with oncologists. Local deployment ensures that model inference occurs entirely within institutional firewalls, thereby reducing privacy risks and strengthening compliance with regulatory standards. Equally important, clinician input is embedded at every stage of development, from prompt design for report restructuring to validation of multimodal prognostic models, so that tools align with established reading patterns and reporting standards. This emphasis on workflow integration and privacy preservation guides our strategy for building AI systems that are both technically advanced and operationally feasible.

To contextualize these efforts, we first review the foundations and current trends in clinical AI for radiology, highlighting advances in language models, computer vision, and multimodal vision–language systems, followed by developments in federated learning, considerations for clinical readiness and translation, and practical deployment pathways (Section 2). We then present Moffitt Cancer Center’s pragmatic approach to integrating AI into oncologic radiology, highlighting real-world deployment and algorithmic development across structured reporting, multimodal cachexia prediction, federated learning, and PHI/PII-redaction workflows (Section 3). Finally, we discuss emerging directions in radiology AI, including tumor board decision support, clinical trial matching, radiology report quality assurance, and imaging complexity index development (Section 4). Our aim is to offer actionable insights for radiology departments seeking to responsibly adopt and integrate AI technologies to enhance diagnostic quality, workflow efficiency, and patient care. The overall organization of the paper is summarized in Figure 1.



**Figure 1. Overview of the paper.** The manuscript is structured around three primary components: (1) foundations and current trends in radiology AI, (2) applied use cases from Moffitt Cancer Center, and (3) emerging directions in radiology AI.

## 2. Foundations & Trends in Radiology AI

AI in radiology encompasses a range of models designed to process medical imaging data, clinical text, or multimodal combinations of both [23–25]. To support collaboration across institutions while preserving patient privacy, federated learning frameworks enable joint model training without requiring data transfer [26]. Bringing these systems into practice requires careful attention to clinical readiness, encompassing model explainability and trustworthiness, radiologist involvement in the loop, and comprehensive strategies to address bias, fairness, and robustness. Once models reach sufficient maturity, they can be deployed through infrastructures that safeguard data security, minimize technical overhead, and integrate seamlessly into clinical workflows, whether maintained locally or supported externally.

### 2.1. Language Models in Radiology

Text is central to radiology, with radiology reports serving as the primary means of documenting imaging findings, diagnostic impressions, and clinical recommendations. In addition to radiology reports, clinical data is also embedded within the electronic health record (EHR) system, which include clinical notes, pathology reports, lab results, medication lists, and imaging summaries. Recent progress in natural language processing (NLP) has enabled new tools for analyzing and working with textual radiology and diagnosis-related data. Central to this progress is the introduction of transformer-based models, which depart from sequential text processing by analyzing a large set of words simultaneously. Through a self-attention mechanism, transformers capture contextual relationships across entire passages, enabling deeper and more flexible language understanding [27]. Early transformer models such as Bidirectional Encoder Representations from Transformers (BERT) laid the groundwork for modern NLP by introducing bidirectional context processing [28].

Building on this foundation, LLMs have emerged as valuable tools for radiology-specific NLP tasks [29]. These models are trained on vast text corpora and contain billions to hundreds of billions of parameters, enabling them to capture far more complex language patterns and nuances. These models can be fine-tuned on specialized datasets, such as annotated radiology reports, to further improve their performance on targeted clinical tasks [30]. In addition to fine-tuning, prompt-based approaches have become increasingly popular. Rather than modifying the model itself, prompt-based methods guide the model's behavior through carefully crafted input instructions [31]. This allows general-purpose LLMs to perform domain-specific tasks with minimal setup, offering flexibility and enabling quick adaptation to new tasks. In radiology, these models are increasingly used for structuring reports, extracting clinical information, and question-answering.

### 2.1.1. Structured Reporting

Structured reporting is one of the examples how NLP can support radiology practice. Studies have shown that structured reports offer greater clarity and completeness compared to traditional free-text reports [32]. Recently, prompting LLMs have been explored as tools for converting unstructured radiology reports into structured formats [33]. Models such as Generative Pre-trained Transformer 4 (GPT-4) and Mixtral have been applied for this task, offering a more flexible and scalable alternative to rigid template-based systems, particularly when combined with domain-specific prompt engineering [3,34].

### 2.1.2. Information Extraction

Information extraction is another essential NLP task in radiology, aimed at identifying clinical entities such as anatomical structures, imaging findings, and diagnostic impressions, as well as markers of uncertainty or temporal information. Extracting these elements accurately is crucial for downstream applications like clinical decision support, research analytics, and integration with structured EHRs. Traditionally, transformer-based models such as BERT have been widely used for this task, often with domain-specific fine-tuning to enhance performance on radiology texts [35]. More recently, prompt-based approaches using open-source LLMs like Vicuna have shown strong results in extracting structured information from unstructured radiology reports, leveraging well-crafted prompts to reduce the need for additional training [4].

### 2.1.3. Question Answering (QA)

QA is an emerging area in radiology where language models are used to respond to natural language queries. This is particularly valuable in clinical settings where rapid access to specific findings is needed. For instance, a model might answer questions like "Is there any evidence of pleural effusion?" or "Has this thyroid cancer progressed?" [36]. Besides radiology reports, questions may draw on broader content from EHRs. Clinical LLM such as GatorTron has demonstrated strong performance in clinical QA tasks involving EHRs, although it was not specifically trained on radiology data [37]. Beyond clinical use, LLMs are increasingly explored for patient-facing applications. When integrated into patient portals, they can help interpret radiology results, explain terminology, and clarify follow-up instructions [38]. LLM-based chatbots, including ChatGPT, Copilot (formerly Bing), and Gemini, have demonstrated high accuracy in responding to patient medical questions, supporting their role in patient education and communication [38–40]. However, unmediated patient-facing LLM-based QA raises concerns related to misinterpretation of indeterminate imaging findings and loss of clinical context, underscoring the need for clinician-supervised deployment [41].

## 2.2. Computer Vision Models for Radiological Imaging

A substantial portion of clinical AI in radiology focuses on the interpretation of medical images, such as computed tomography (CT), magnetic resonance imaging (MRI), positron emission tomography-computed tomography (PET-CT), and radiographic "X-ray" scans. Image-based models used in this context generally fall into three categories: convolutional neural networks (CNNs), vision transformers (ViTs), and hybrid architectures that combine the two [42–44].

CNNs are the most widely used in radiological visual recognition tasks. These models work by scanning images with small filters to detect local patterns such as edges, shapes, and textures [45]. ViTs are a newer class of models that divide an image into patches, analogous to tokens in text, and process these patches in parallel while learning the relationships between them [46]. This allows ViTs to capture fine-grained spatial detail along with broader contextual cues, making them well-suited for radiology tasks where subtle, distributed patterns are often diagnostically important [47,48]. Unlike CNNs, which process data hierarchically, transformers evaluate spatial relationships more holistically. However, ViTs typically require larger datasets and greater computational resources, which has limited their widespread adoption in clinical workflows thus far [46,48]. Recent approaches also begun integrating CNNs with transformers, leveraging the strengths of both: the local feature extraction of

CNNs and the global contextual understanding of transformers. In radiology, image-based models are often used for segmentation, detection, and classification tasks [49].

### 2.2.1. Segmentation

Segmentation involves outlining regions-of-interest including organs, nodules, tumors, or edema. When done by hand, this process can be time-consuming and may vary between doctors. AI-based methods help by making segmentation faster, more consistent, and less dependent on individual judgment [50]. One of the most widely used architectures for medical image segmentation is U-Net, a CNN-based model with an encoder–decoder structure and skip connections to preserve spatial resolution [6]. Originally designed for 2D image segmentation, U-Net processes individual slices and is often applied to X-rays, ultrasound, or single CT/MRI slices [6,51–53]. For volumetric data, 3D U-Net extends this architecture to entire image volumes, capturing spatial context across slices, especially useful for CT and MRI [42]. Numerous other variants of U-Net have been developed to improve performance in radiological applications [54,55]. Despite their success, CNN-based models can struggle with modeling long-range spatial dependencies in high-resolution images.

ViTs have been explored as a solution to this limitation [43,56]. Moreover, hybrid architectures integrate ViTs with CNNs to benefit from both local and global information [47]. A common approach is to incorporate ViTs into different segments of the U-Net architecture. For instance, ViTs can be integrated into the encoder to enhance feature representation [48], into the decoder to improve the reconstruction of segmentation maps from encoded features [57], or into the skip connections to enable richer fusion of local and global features across network levels [58].

### 2.2.2. Detection

Object detection refers to identifying the presence and location of abnormalities, such as lung nodules, brain lesions, or bone fractures. This typically involves highlighting suspicious regions with bounding boxes and classifying their type. CNN-based object detection models are widely used in medical image detection, and are typically classified as two-stage or single-stage models [59]. Two-stage models first propose regions of interest and then classify them [60]. Single-stage models perform detection and classification in one pass, offering greater speed and efficiency [7,61]. In addition to 2D CNNs, 3D CNNs have been developed to handle volumetric medical data more effectively [62]. More recently, ViTs have been explored for medical image detection, primarily for 2D tasks, due to their ability to capture long-range dependencies. However, their reliance on large datasets and high computational resources has limited their standalone use. Instead, ViTs are often combined with CNNs in hybrid architectures, where they can be integrated into the backbone, neck, or head of the network to enhance feature representation and improve detection performance [63].

### 2.2.3. Classification

Image classification is the task of assigning diagnostic labels to images or localized regions, such as distinguishing between benign and malignant findings or identifying tumor subtypes. CNNs (e.g., ResNet, DenseNet) are the backbone of most classification systems due to their proven ability to detect patterns in both 2D and 3D imaging data [8,64]. ViTs have also been used for classification tasks, showing stronger performance than CNNs [65,66]. Combining ViTs with CNNs has become increasingly popular as well [67]. Some models, such as TransMed, extend image-based classification by incorporating multiple MRI modalities, specifically T1-weighted and T2-weighted scans, using a hybrid CNN–Transformer architecture [44]. Transfer learning is widely used in this domain, particularly with CNNs, where models pre-trained on large natural image datasets like ImageNet are fine-tuned on smaller medical datasets [68]. This approach significantly improves performance, especially when labeled data is limited. More recently, transfer learning has also been applied to ViTs, though it remains less common due to the larger data and compute requirements for effective pretraining [65]. Radiology classification tasks often involve multi-label outputs since a single scan can reveal multiple co-existing pathologies which requires models that can handle overlapping labels [64].

### 2.3. Vision-Language Models (VLMs) & Multimodal AI in Radiology

Recent advances in AI have led to the development of VLMs that process both visual and textual data [14]. In radiology, this often involves combining visual information from imaging with associated clinical text such as pathology reports, patient history, or lab results. By jointly reasoning over both text and image data, these models provide richer contextual understanding than either modality alone [25].

VLMs are typically organized into single-stream and dual-stream architecture designs [14]. Single-stream models process both image and text data jointly within a unified transformer architecture, allowing for deep interaction between modalities from the early layers. These models often incorporate a vision encoder, such as a CNN or ViT, followed by a projection module that maps the visual features into the same feature (or embedding) space as text tokens. The projected visual features are then combined with tokenized text inputs and jointly fed into an LLM [69,70]. Dual-stream models use separate encoders for image and text inputs without sharing parameters. For example, images may be encoded using a CNN or ViT, while text is processed using an LLM. The modality-specific representations are then aligned at later stages, often using attention-based fusion layers [71]. In models that generate free-text outputs, such as radiology reports or answers to clinical questions, a decoder module is typically introduced in the end [69,71]. In radiology, VLMs are commonly applied to tasks such as report generation, VQA, and prognosis prediction.

#### 2.3.1. Report Generation

Radiology report generation involves identifying key findings, assessing their clinical significance, and clearly communicating the results of imaging studies, such as X-rays, CT scans, or MRIs, to referring physicians. This task is time-consuming and subject to variability in style and completeness. Leveraging VLMs can streamline this workflow by automatically generating structured reports directly from medical images [14]. Most existing VLMs focus on chest X-rays, primarily due to the availability of large paired image-text datasets such as MIMIC-CXR [72]. Recent models including XrayGPT [73], RaDialog [70], RadFM [16], Flamingo-CXR [9], and RadAlign [12] have been fine-tuned on MIMIC-CXR to capture radiology-specific language and image features.

In contrast, there has been comparatively less work on generating radiology reports from 3D radiology images, such as CT and MRI. Med3D-VLM [11] and Med-Gemini-3D [15] integrate 3D medical image encoders to support volumetric modalities, while RadAlign [12] and RadVLM can handle multiple imaging types, including X-ray, CT, and MRI. In report generation, models often utilize task-specific prompts to guide their output toward a desired structure, thereby improving clarity and completeness. Performance can be further enhanced by leveraging CheXpert [74] labels. For example, through a double-feature transformer that integrates a pretrained CNN encoder, such as ImageNet, with a chest radiograph-specific CNN encoder trained on CheXpert labels [75]. More recently, retrieval-augmented generation (RAG) has been incorporated to retrieve relevant, clinically validated report segments from large datasets, resulting in reduced hallucinations and improved factual grounding [12,76].

#### 2.3.2. Visual Question Answering (VQA)

In radiology, VQA requires the model to interpret an image together with a question to produce an accurate and relevant answer. Queries can range from identifying abnormalities or medical conditions to specifying the imaging modality, anatomical region, or presence of certain structures. They are generally open-ended, allowing detailed free-text responses, or closed-ended, with predefined answers such as "yes" or "no," multiple-choice options, or numeric values. VQA can be approached as a classification task, selecting from a fixed set of answers, or as a generation task, producing responses in free form [14]. Models are frequently fine-tuned on datasets such as VQA-RAD [77] and SLAKE [78], which include X-rays, CTs, and MRIs. Recent approaches also integrate RAG to improve factual accuracy and reduce hallucinations by grounding answers in clinically verified text corpora [79].

Conversational VQA extends standard VQA by enabling multi-turn, context-aware interactions in which follow-up questions can refer back to prior questions or answers. This allows radiologists or clinicians to conduct a dialogue with the AI system, for example, first asking whether a lesion is present, then requesting its size or location without having to restate the full context. Such systems require maintaining conversational memory and reasoning across turns, often integrating dialogue management components on top of standard VQA pipelines. The most recent VLMs that support this task include RaDialog [70] and RadVLM [10].

### 2.3.3. Prognosis Prediction

Prognosis prediction involves estimating a patient's future clinical outcome or the likely course of a disease, using imaging, often in combination with other clinical or demographic information. Typical objectives include forecasting survival duration, assessing the probability of disease progression, or determining the risk of recurrence. One recent multimodal approach for pulmonary embolism survival estimation integrates CT pulmonary angiography (CTPA) imaging with automatically generated report segments from a VLM, abnormality classification outputs, and Pulmonary Embolism Severity Index (PESI) scores, with the combined data analyzed using a Cox regression framework [80]. Another example is a three-dimensional FM for abdominal CT that jointly processes volumetric CT data, corresponding radiology reports, and structured diagnosis codes from EHRs. This system supports a range of downstream applications, including the prediction of chronic disease occurrence over a five-year period [81].

### 2.4. Federated Learning: Transforming Radiology AI Collaboration

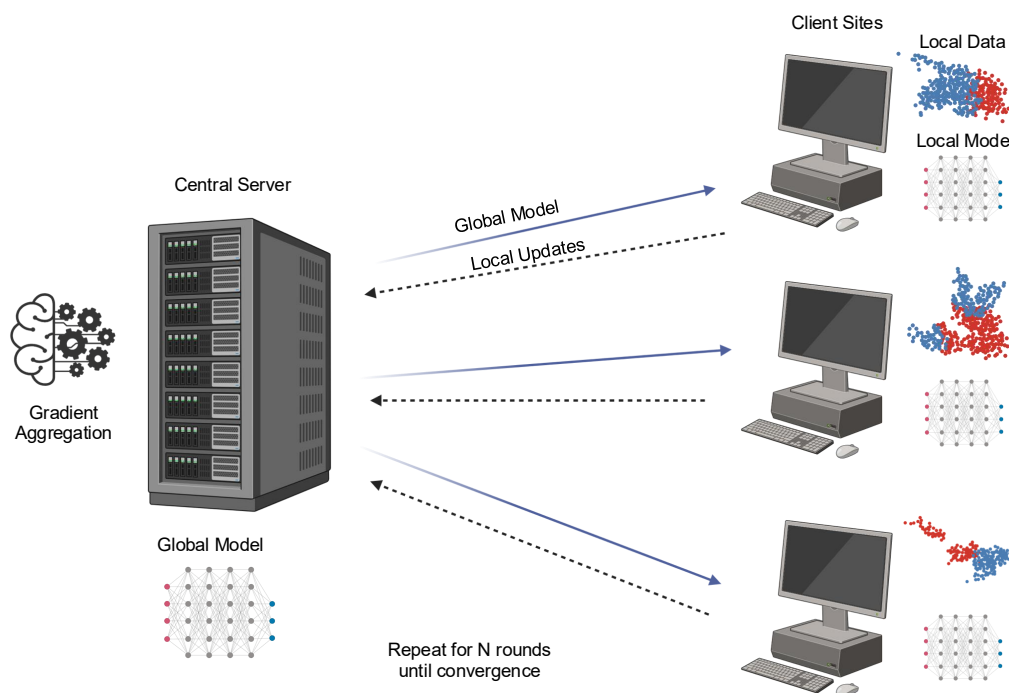
Training accurate and generalizable radiology models requires access to very large datasets. Most individual institutions have limited patient volumes and do not capture the full diversity of cancer types, stages, and clinical presentations needed for robust oncology research [82]. Privacy regulations such as the Health Insurance Portability and Accountability Act (HIPAA) in the United States and the General Data Protection Regulation (GDPR) in Europe place strict conditions on the transfer of PHI [83,84]. These regulations do not prohibit collaboration, but they require strong safeguards and formal approvals for any movement of patient data. Because these requirements can be complex and time consuming to satisfy, many institutions are unable or unwilling to send patient data to a shared central location, which makes centralized research datasets difficult to assemble at large scale.

FL offers a practical way to enable collaboration under these constraints [26]. In FL framework, each institution keeps its data within its own secure environment and trains a local version of the model on its own dataset. After one round of training, the institution sends only the updated model parameters to a coordinating server. The server combines these updates to form a global model and then returns the global parameters to all participating sites. This process repeats for many rounds and is shown in Figure 2. The approach was first formalized in the Federated Averaging (FedAvg) method, in which the global model is produced by computing a weighted average of the local updates [85].

FL has been applied in research across many radiology tasks [26,86]. The Federated Tumor Segmentation effort, which developed from the BraTS initiative, is the largest medical federated project, with participating sites on six continents [87]. FL also enables multimodal development, allowing imaging, clinical notes, laboratory values, and electronic health record information to be used collectively while all data remain within the institution where they were generated [88].

Although FL avoids direct data sharing, it still faces privacy and security challenges [89]. Local model updates can, in some cases, leak information about the underlying data, and studies have shown that sensitive details can sometimes be inferred through attacks that attempt to reconstruct patient information from gradients [89,90]. Modern FL frameworks include several safeguards to reduce these risks. Secure aggregation protocols prevent reconstruction of local updates during transmission [91]. Differential privacy adds carefully controlled noise to the model parameters to reduce the chance of patient re-identification [92,93]. Homomorphic encryption allows calculations to be performed on encrypted numeric values, maintaining confidentiality throughout the process [94,95]. Taken together,

these features make FL a promising path toward large scale collaborative radiology research that preserves both data utility and patient privacy.



**Figure 2.** Overview of the FL process. Each institution trains a local model on its own data and sends only updated model parameters to a coordinating server, which produces a global model by aggregating the local updates.

## 2.5. Clinical Readiness & Translation

Despite rapid advancements in the capabilities of AI models, successful integration into radiology clinical practice requires ensuring clinical readiness and effectively translating these capabilities into practical applications.

### 2.5.1. Model Explainability & Trustworthiness

For radiology AI models to be clinically adopted, they must produce outputs that radiologists can interpret and verify. Black-box predictions undermine trust, particularly in high-stakes diagnostic contexts such as oncology. A range of explainability techniques can be applied, depending on the model type, whether it is language-based, vision-based, or a multimodal VLM.

For language-based radiology AI, attention weight visualization is a common approach for identifying which tokens a model attends to during prediction. However, its validity is not fully established, as attention does not always align with the truly influential features [96]. One alternative is TokenSHAP, an adaptation of SHapley Additive exPlanations (SHAP) [97] tailored to sequential data [98]. TokenSHAP assigns each token an importance value based on cooperative game theory, while preserving the dependencies between words, and offers a more reliable measure of token influence than raw attention scores alone. Building on attribution-based insights, chain-of-thought reasoning offers a complementary perspective by revealing the model's intermediate reasoning steps, either extracted from hidden states or generated explicitly [99]. Recent approaches enhance this by adding an explicit reasoning component, where an LLM generates not only its final output but also a human-readable explanation grounded in evidence from the input [100]. This reasoning trace can then be audited for factual accuracy, increasing trust in AI-assisted decision-making.

In computer vision models, guided backpropagation examines how a model's output changes when parts of the input image are perturbed, helping to identify which pixels most influence the prediction [101]. Techniques based on the class activation mapping (CAM) [102], such as Grad-CAM [103] and Score-CAM [104], localize the image regions that contribute most to a model's decision,

providing a visual heatmap of the most important features. Concept Activation Vectors (CAVs) measure how strongly a clinically relevant concept (e.g., a lung nodule) influences a model's predictions across many cases [105]. Unlike previous methods, which explain a single prediction, CAVs enable the global testing of concepts, making them valuable for detecting hidden biases and verifying that the model bases its decisions on medically meaningful features rather than spurious patterns. In radiology, for example, Grad-CAM has been used to highlight critical areas in brain MRIs [106], Score-CAM has been applied to mammography images [107], and CAVs have been employed in chest X-ray analysis [108].

In the context of VLMs for radiology, newer adaptations extend these ideas to align image regions with generated text. VL-InterpreT is an interactive visualization system designed for multimodal transformers [109]. It enables the examination of both intra- and cross-modal attention patterns and tracks how hidden representations evolve across layers for both image and text tokens. MM-SHAP [110] adapts the SHAP framework to multi-modal models, assigning importance scores to both text tokens and image patches. However, its reliance on patch-level granularity, rather than semantically meaningful objects, limits the clinical interpretability of its visual outputs. PixelSHAP [111] builds on this approach for text-generative VLMs by attributing importance to structured visual entities (e.g., anatomical regions), producing explanations that are more aligned with human visual reasoning.

### 2.5.2. Human-in-the-Loop Oversight

Human-in-the-loop (HITL) oversight refers to the continued involvement of qualified professionals in the operation of an AI system, ensuring that model outputs are reviewed and corrected when necessary before they influence patient care [112]. This is particularly important in radiology, where diagnostic decisions carry high clinical stakes and imaging findings vary widely across patient populations. HITL ensures that responsibility for the final interpretation remains with the radiologist, positioning AI as an assistive rather than an autonomous agent.

HITL engagement occurs across multiple stages of the AI life cycle. During development, radiologists may annotate datasets, validate algorithmic outputs, and provide structured feedback that guides the refinement of the model. Active learning approaches exemplify this stage: by algorithmically identifying high-yield or uncertain cases for expert labeling, these methods reduce annotation burden while ensuring that physician expertise is directed where it is most needed [113]. Reinforcement learning with human feedback (RLHF) has also been proposed as a complementary strategy, where expert evaluations are integrated into training to align models with clinically meaningful outcomes [114].

In deployment, HITL most often takes the form of radiologists reviewing and editing AI-generated outputs before they are entered into the official patient record. This preserves diagnostic authority while enabling workflow gains. A recent study showed that radiologists using AI-assisted draft reports achieved a reduction in average reporting time from 573 to 435 seconds without an increase in clinically significant errors [115]. Gaze-tracking-based oversight has also been proposed, where eye-movement data highlight regions that may have been overlooked by either the AI or the radiologist, thereby reducing perceptual misses [116]. In some deployments, clinician adjustments to AI outputs are recorded and later reviewed in local governance meetings, providing a structured feedback mechanism for oversight [117]. Other work has demonstrated frameworks where AI results are delivered through DICOM structured reporting, allowing radiologists to directly validate or reject outputs in their reading environment and ensuring a transparent feedback loop between the system and its users [118].

### 2.5.3. Robustness, Bias, & Fairness in Radiology AI

Many radiology AI models perform well on retrospective or single-institution datasets but often lose accuracy when applied across broader clinical populations. In a review of 83 published radiology AI studies, almost half of the models that underwent external validation showed at least a moderate drop in accuracy, and nearly one-quarter experienced a major decline compared with their internal test results [119]. For example, the ICOVAI system for COVID-19 CT and several mammography models both showed strong internal performance but declined when tested on external cohorts [120,121]. Another model trained for pneumonia detection relied partly on hospital-specific image characteristics,

leading to reduced diagnostic reliability at external sites [122]. These findings illustrate the need for systematic testing across diverse sites before clinical adoption.

Performance differences are not only institutional but also demographic. For instance, one study showed that models trained on male-dominant datasets demonstrated reduced accuracy in female patients, creating systematic bias in chest pathology detection [123]. Another work has demonstrated that race can be inferred from medical images even though it is not visible to human readers, raising the concern that such hidden signals may influence predictions in unintended ways [124].

Several approaches are being developed to address these challenges. FL allows collaborative training of models across institutions without sharing raw data (e.g., DICOM images), and it can achieve performance that is comparable, and even superior, to centralized models while maintaining privacy and institutional control [125,126]. The federated models, owing to the way they are trained, may result in more generalizable models addressing the challenges of models trained using a single side data. Domain adaptation seeks to harmonize imaging data from different scanners or acquisition protocols so that models trained in one environment remain reliable in another. For example, attention mechanisms can guide models toward features that remain consistent across sites, while disentangled networks can harmonize scans to appear more comparable without retraining [127,128]. Data augmentation also plays a role in improving generalization and addressing data imbalance. Diffusion models generate high fidelity synthetic medical images by learning to reverse a gradual noising process, allowing realistic creation of CT and MRI examples that expand limited datasets and support more robust model performance across varied clinical settings [129]. To reduce demographic bias, techniques such as REPAIR reweight training samples to encourage learning of features that generalize beyond overrepresented groups [130]. Adversarial debiasing takes a different approach, using an auxiliary network to detect sensitive attributes and penalizing the predictor when those attributes are encoded, producing outputs that are less dependent on demographic factors [131].

## 2.6. Deployment Pathways

Effective radiology-AI deployment depends on where models run, how securely and reliably they are operated, and how seamlessly they fit in the clinical workflow.

### 2.6.1. Local vs. Cloud-Based Deployment of AI Models

Hospitals that want to use AI in radiology face an important choice: deploy the AI models on local computation hardware or connect to the provider's remote servers (i.e., cloud-based deployment). Local deployment means the model runs on servers that are physically available inside the hospital's network. For example, platforms such as Ollama and vLLM allow LLMs/VLMs to be downloaded and run locally [34]. Local systems may have lower latency because they avoid delays introduced from sending data over the internet. However, local deployment requires institutions to invest in computer equipment such as graphics processing units (GPUs) and to maintain software updates themselves.

Cloud-based services work differently. Data are sent over secure internet connections to a vendor's servers, where the AI model are deployed. The results are then sent back to the institution. This arrangement reduces the technical burden for the institutions and ensures they are always using the latest model version. Well-known examples include Viz.ai, RapidAI, and Aidoc, which provide FDA-cleared cloud-based radiology AI applications for stroke triage, pulmonary embolism detection, and other urgent findings. General-purpose platforms such as Amazon Web Services (AWS) also provide the infrastructure to host radiology AI at scale, and many vendors build their services on top of these environments. The trade-off is that data leave the institution, so hospitals must rely on strong encryption, legal agreements, and regulatory compliance to protect PHI. In most cases, these services are accessed through application programming interfaces (APIs), which act as standardized gateways for sending data to the vendor and receiving model outputs in return. APIs make integration with existing systems easier but also create another layer of dependency on the vendor's infrastructure and uptime.

To simplify both approaches, many developers now deliver their software in containers (e.g., Docker). A container packages the AI model together with all the source code and dependence libraries it needs to run, so the same container can be installed locally behind the firewall or deployed on cloud servers by the vendor. In radiology AI, a leading example is the MONAI Deploy Toolkit, which is designed to streamline the packaging, validation, and integration of AI applications into clinical imaging environments [132]. Using this framework, AI models can be packaged into containers that connect directly with institutional systems, including the Picture Archiving and Communication System (PACS) and Radiology Information System (RIS). This setup allows imaging studies to be processed securely and re-integrated into clinical workflows within the institutional firewall.

### 2.6.2. Secure AI Infrastructures

The deployment of AI in radiology requires more than deciding between local or cloud-based hosting. It also depends on secure infrastructures that can protect imaging data, maintain reliable performance, and integrate safely into clinical workflows. Because radiology data contain sensitive personal and diagnostic information, breaches or failures can have significant clinical and legal consequences.

A core requirement is cybersecurity. Imaging networks have been frequent targets of ransomware, and vulnerabilities in PACS and hospital IT systems have been exploited in the past [133]. Effective safeguards include proactive vulnerability assessments, strong authentication, encryption of imaging data in transit and at rest, and formal incident response protocols [134]. These protections are especially critical when AI tools integrate with PACS, RIS, or EHRs, where unauthorized access could expose both imaging data and clinical metadata.

Equally important is lifecycle monitoring of deployed models. Unlike static software, radiology AI can lose accuracy over time as scanners are upgraded, imaging protocols evolve, or patient populations shift. Without oversight, this “model drift” can reduce clinical reliability. Continuous monitoring frameworks have been proposed in radiology to track performance in real-world use, flag model decline, and trigger model recalibration or retraining when needed. Some authors describe this as the move toward continuous learning AI, which brings both opportunities for adaptation and new requirements for governance and validation [135].

Finally, institutional governance plays a central role. Many institutions are beginning to establish AI oversight committees or integrate AI risk management into existing quality assurance programs [117]. These groups review proposals for new AI deployments, set guidelines for validation, and monitor safety once systems are in use. Such governance ensures that radiology AI does not function as a “black box” add-on but as a monitored tool within the broader framework of patient safety and institutional accountability.

### 2.6.3. Integration with Clinical Workflow

Beyond the decision to host applications locally or in the cloud, and aside from the technical safeguards that protect medical data, the long-term value of radiology software depends on whether it fits smoothly into day-to-day workflow. Radiologists are far more likely to use decision-support tools when results appear directly in the software programs they already depend on, such as PACS viewers, structured reporting systems, or electronic templates. Research has shown that AI applications must be presented with minimal extra steps and implemented in a way that does not disrupt existing reporting routines; otherwise, they are unlikely to be embraced in daily practice [136].

One way hospitals are addressing this problem is by adopting vendor-neutral platforms. For example, the ROCKET platform, which routes studies to selected algorithms and returns the findings into PACS using DICOM Structured Reporting. This setup allows radiologists to review results in the same environment where they interpret images and also provides mechanisms to accept, reject, or request revisions, ensuring that AI remains embedded within the reporting workflow [137].

Institutions are also beginning to adopt vendor-neutral AI (VNAI) platforms to simplify integration at scale. These platforms serve as an intermediary between institutional imaging infrastructure

and a range of AI applications. They handle data routing, standardize interfaces, and return outputs directly into PACS or reporting systems. By consolidating connections, VNAI platforms reduce technical overhead and give institutions the flexibility to introduce or retire applications without being locked to a single supplier. This architecture makes it possible to support multiple clinical use cases while keeping results embedded in the radiologist's familiar workflow [138].

### 3. Use Cases from Moffitt Cancer Center

#### 3.1. Structured Radiology Reporting Using Local LLMs

Radiology reporting plays an essential role in cancer care by guiding diagnostic interpretation, staging, and treatment planning. However, radiology reports can often be lengthy, formatted in various ways, and variable between individual radiologists [139]. This lack of standardization poses challenges for referring oncologists, who may be interested in quickly extracting key diagnostic information to inform patient management. Furthermore, equally important is the fact that these radiology reports will serve as a data point for future AI model training. Consistently structured reports are an institutional asset for future AI development. At Moffitt Cancer Center, we recognized the need for a scalable and privacy-preserving solution to improve the clarity, conciseness, and structural consistency of radiology reports, particularly in the context of high-volume oncology imaging.

To address this need, we developed and implemented an LLM-based radiology report processing pipeline aimed at restructuring and streamlining radiology reports [34]. Central to our approach was a commitment to safeguarding patient privacy by ensuring that all model inference occurred locally, behind our institutional firewall. This local deployment strategy allowed us to avoid reliance on external cloud-based APIs, thus eliminating the risks associated with transmitting PHI outside institutional control.

Our technical workflow involved evaluating several open-weight state-of-the-art LLMs, such as Mixtral [140], Mistral [141], and Llama 3 [142]. All models were deployed on institution-owned workstations equipped with standard hardware configurations, including RTX 3060 GPUs with 12 GB of VRAM. The Ollama framework facilitated efficient local inference, and LangChain was used to implement flexible prompt engineering workflows tailored to radiology-specific tasks.

A key component of our development process involved optimizing prompt strategies to balance two primary goals: reducing report verbosity and ensuring its structural consistency. Working closely with board-certified body radiologists, we designed and tested five distinct prompting strategies. Among these, the two-step *Conciseness then Structure (C » S)* strategy consistently produced the most reliable results. This approach first condensed the original report to eliminate redundancy and unnecessary phrases, followed by a second step that reorganized the content into a standardized, organ-based template. We also incorporated an automated output validation step using an *OutputFixingParser*, which prompted the LLM to correct its formatting if initial outputs did not follow the desired structure.

In a retrospective quality improvement study, we applied the LLM pipeline to 814 de-identified radiology reports covering CT scans of the chest, abdomen, and pelvis from cancer patients. These reports were authored by seven board-certified radiologists from our center. The results demonstrated a substantial reduction in report length, with an average word count decrease exceeding 53% across all reports. Beyond improving conciseness, the LLM-processed reports exhibited greater structural uniformity, potentially making it easier for referring oncologists to locate pertinent findings. Preliminary feedback from radiologists underscored the perceived benefits of the tool in enhancing report readability and communication with clinical teams.

Building on the success of these initial results, we have recently expanded our LLM-based radiology report pipeline to incorporate reasoning-focused LLM capabilities. A recognized limitation of current LLM outputs is the absence of transparent, interpretable reasoning behind the generated content. To address this, we now employ reasoning-enabled models, including, DeepSeek [100] and QwQ [143]. These LLMs extract and present intermediate reasoning steps during report transformation to the user for transparency and validation. These reasoning traces include justifications for why certain

findings were categorized under specific organ systems, explanations for the removal of redundant phrases, and confidence indicators for each section of the output. By exposing this internal decision-making process, we aim to establish a more interactive AI-human feedback loop. Currently, our radiology team is conducting a review of both the final structured reports and the underlying reasoning pathways, to provide feedback for informing further model refinement and prompt optimization.

Through this iterative approach, our team is moving toward a more explainable and collaborative paradigm of AI deployment in radiology. The lessons learned from this use case, particularly around local deployment feasibility, prompt strategy design, and clinician-AI interaction, are informing our broader institutional strategy for integrating AI tools into cancer imaging workflows. Most importantly, this initiative underscores the potential for LLMs to deliver practical, workflow-compatible solutions that enhance cancer diagnosis and treatment planning.

### 3.2. *Imaging-Informed AI for Early Cancer Cachexia Diagnosis*

Cancer cachexia, a multifactorial metabolic syndrome characterized by substantial weight loss, skeletal muscle wasting, and systemic inflammation, remains one of the underdiagnosed oncology conditions despite its profound impact on patient survival and quality of life. Up to 80% of cancer patients are impacted by cachexia with high prevalence in certain cancers, such as gastroesophageal, pancreatic, colorectal, lung, and hematological [144–146]. Existing diagnostic mechanisms are often limited by manual processes, fixed thresholds, and poor integration into clinical workflows [147–149]. At Moffitt Cancer Center, we have developed clinically deployable AI tools to bridge this gap, focusing on reliability, interpretability, and workflow compatibility to bring imaging intelligence directly into patient care.

Skeletal Muscle Assessment–Automated and Reliable Tool based on AI (SMAART-AI) introduces an uncertainty-aware imaging data processing framework for automated skeletal muscle segmentation from routine CT scans. Rather than optimizing for accuracy alone, the SMAART-AI prioritizes reliability and trustworthiness, a critical step in clinical deployment [150,151]. Unlike existing commercial segmentation tools that may silently degrade under input data domain shifts (e.g., differences in scanner type, protocol, image quality, population demographics, or events like a pandemic), SMAART-AI quantifies prediction uncertainty and flags potentially unreliable outputs, enabling HITL validation. Training strategies and reliability-based design make the tool robust in heterogeneous cancer cohorts. SMAART-AI also supports longitudinal tracking of skeletal muscle area (SMA) and skeletal muscle index (SMI), providing insight into changes throughout the treatment journey. The imaging-derived metrics are further integrated with structured and unstructured patient data for downstream tasks such as cachexia risk and survival prediction.

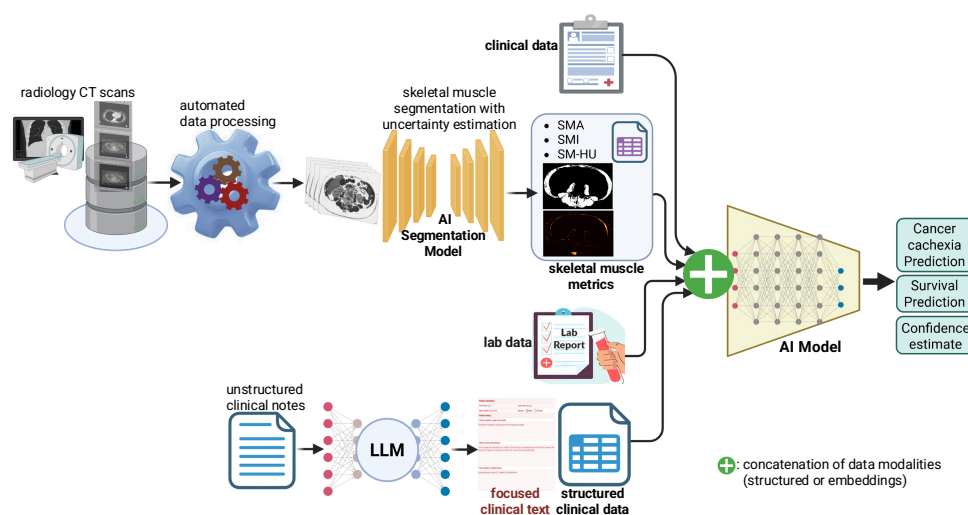
Building on SMAART-AI, our team developed a “Multimodal AI Biomarker” framework that extends radiologic insight into a broader, data-driven pipeline for early cachexia diagnosis. The Multimodal AI Biomarker fuses CT-derived skeletal muscle metrics with structured clinical variables (e.g., age, sex, race, ethnicity, weight, height, BMI, and stage of disease), blood-based laboratory biomarkers (e.g., serum albumin, neutrophil, lymphocyte count, blood urea nitrogen, creatinine, and derived ratios such as neutrophil-to-lymphocyte and blood-urea-to-creatinine), and cachexia-related symptoms extracted from unstructured clinical notes using locally-deployed LLMs. The framework adapts dynamically to patient-specific factors and missing data modalities, ensuring compatibility with real-world clinical data and readiness for potential clinical deployment.

In a retrospective evaluation of 236 patients with pancreatic cancer from the Florida Pancreas Collaborative cohort, the step-wise addition of each data modality (imaging, clinical, lab, and physician’s notes) improved model performance, achieving 77% accuracy using clinical and imaging data, and 85% accuracy when laboratory and LLM-extracted symptom features were combined [152]. In a separate embedding-based implementation, CT image embeddings corresponding to the L3 level generated using the RadImageNet FM [153], were fused with embeddings from GatorTron [37] for clinical, laboratory, and free-text medical notes data, enabling joint multimodal representation learning. Incorporating these radiology-derived image embeddings further enhanced cachexia prediction ac-

curacy to 92%, underscoring the additive value of radiologic information in multimodal fusion [152]. Similarly, survival prediction improved progressively, with concordance index (C-index) gains from 0.64 using clinical data alone to 0.72 when skeletal muscle metrics, laboratory indicators, and clinical text features were added. Removing skeletal muscle metrics led to considerable declines in both cachexia diagnosis accuracy and survival prediction performance, highlighting the indispensable prognostic role of radiology-derived features [152].

From the clinical deployment standpoint, all model inferences and patient data processing, including LLM-based extraction from unstructured medical notes, are performed locally within our secure computing infrastructure, avoiding reliance on external APIs and ensuring compliance with institutional data governance policies.

Together, SMAART-AI and the Multimodal AI Biomarker exemplify how imaging-derived information serves as a cornerstone for multimodal clinical AI, advancing early diagnosis, personalized intervention, and scalable cachexia management. Our ongoing work is focused on expanding this framework to a multi-institutional settings through FL, supporting privacy-preserving model development across oncology centers. These efforts reflect our commitment to developing patient-centered, trustworthy AI systems designed for safe, transparent, and clinically integrated oncology practice.



**Figure 3.** Multimodal AI workflow for cachexia prediction. The framework integrates imaging-derived skeletal muscle metrics generated by SMAART-AI with structured clinical variables, laboratory biomarkers, and cachexia-related symptoms extracted from unstructured physician notes using locally deployed LLMs. Each data modality is processed through dedicated feature extraction modules and fused into a unified multimodal representation for cachexia classification and survival prediction.

### 3.3. Privacy-Preserving and Multi-Institutional AI Collaboration Using FL

At Moffitt Cancer Center, we have implemented and validated a privacy-preserving FL framework to address the challenges of collaborative AI development between institutions while maintaining strict data governance standards. Using the NVIDIA Federated Learning Application Runtime Environment (FLARE) Software Development Kit (SDK), our infrastructure enables distributed training of AI models without the need to share raw patient data. This approach is particularly well-suited for multi-institutional efforts focused on early cancer detection and the development of imaging-based biomarkers.

Our initial FL simulations focused on lung cancer screening, using low-dose CT (LDCT) images from the National Lung Cancer Screening Trial (NLST) [154] and a cohort of real-world patients from the Moffitt Cancer Center. The Moffitt patient data varied significantly in population and characteristics of the two NLST cohorts, providing a real-world testbed to evaluate model generalization under heterogeneous conditions. To simplify onboarding across collaborating institutions, we developed a “One-Touch FL” framework in which each client site connects to the central FL server using a

single Linux terminal command. This approach enables collaborators who may have limited FL experience to participate in FL runs by reducing the setup and configuration burden and mitigating the steep learning curve associated with establishing FL infrastructure. All model development orchestration, including local training, secure model updates, and aggregation, is handled centrally, reducing technical complexity at the partner sites.

Using a standard FedAvg strategy within NVIDIA FLARE, both centralized and federated models were trained on NLST patients to evaluate the generalization gap between them. The FL global model achieved performance comparable to its centralized counterpart, demonstrating the effectiveness of FL for this task. Next, the patient cohort from Moffitt was added as a third site, and a new FL model was trained across all three datasets. This model was then evaluated on a withheld test set of patients from the Moffitt cohort and compared against a centralized model trained solely on the Moffitt patients. The task was binary cancer classification, with the model outputting either “yes” or “no”. Initially, the global FL model performed significantly worse than the Moffitt-only centralized model on the test set, underscoring a clear domain gap between the NLST data and the Moffitt cohort. To address this, the global model was fine-tuned on the Moffitt training data to better align it with the local distribution, a process known as personalized FL. These findings suggest that features learned from the NLST cohorts can be effectively leveraged in the Moffitt domain when the global model is properly adapted to site-specific data.

To evaluate NV-FLARE server-client communication and readiness for model training, we conducted cross-institutional experiments using the CIFAR-10 dataset. Training was performed between Moffitt and other FL sites, using a centrally hosted NV-FLARE server within Moffitt’s AWS Virtual Private Cloud (VPC). This successful deployment confirmed the technical deployment of NV-FLARE for secure asynchronous federated training between geographically distributed sites. Currently, we incorporate multiple privacy preservation strategies, including secure server aggregation and optional integration of differential privacy or homomorphic encryption for enhanced protection. We are integrating several “uncertainty quantification” methods, including ensemble learning and conformal prediction, to enable model confidence estimation in distributed settings. This work is needed to enable HITL workflows in which uncertain predictions could be flagged for clinical review. Current efforts focus on scaling Moffitt’s FL infrastructure to support multimodal model development between institutions that are part of the National Cancer Institute’s Early Detection Research Network (NCI’s EDRN) and the Privacy-Preserving Federated Learning (PPFL) consortium.

#### 3.4. PHI/PII Redaction from Radiology Images and Reports

A critical barrier to secondary use of medical imaging data for AI model development and clinical research is the presence of PHI and Personally Identifiable Information (PII) in both DICOM metadata and pixel data. At Moffitt Cancer Center, in collaboration with Impact Business Information Solutions (IBIS), a novel PHI/PII de-identification framework has been developed that addresses these challenges using a hybrid AI-driven and rule-based approach with integrated uncertainty quantification.

The framework employs a two-tiered pipeline for metadata and pixel data de-identification:

*Metadata De-Identification:* A rule-based system removes explicit PHI from DICOM headers. This is supplemented by a fine-tuned LLM-based Named Entity Recognition (NER) pipeline trained on synthetic clinical data. By simulating PHI instances using synthetic data, we ensure robust identification of sensitive fields without exposing real patient data. Additionally, fuzzy string matching techniques are applied to catch near-variants of detected PHI.

*Pixel Data De-Identification:* For image pixel data, an uncertainty-aware Faster R-CNN model detects and localizes burned-in text regions, a common but under-addressed privacy risk in medical imaging. Detected text regions undergo Optical Character Recognition (OCR) step, and the extracted text is processed through the same NER pipeline for PHI identification and removal. The model integrates uncertainty quantification, enabling the system to assess the confidence of each detection and flag uncertain cases for HITL review.

The AI-assisted hybrid workflow enhances both scalability and reliability. The use of uncertainty metrics provides transparency and allows for HITL verification when the model confidence is low. The de-identification solution has been benchmarked against regulatory standards including HIPAA, GDPR, and TCIA's best practice guidelines. Across over 580,000 data elements evaluated, the system achieved over 99.8% de-identification accuracy, including successful removal of embedded PHI from diverse imaging modalities (CT, MRI, and X-ray). The combination of automation and uncertainty-aware HITL review supports scalable, secure sharing of medical imaging datasets for AI development without compromising patient privacy.

A key innovation of this framework is its explicit risk calibration. Uncertain predictions are transparently quantified, enabling operational teams to set confidence thresholds, quarantine ambiguous data, and meet specific institutional or jurisdictional privacy requirements (such as "Expert Determination" under HIPAA).

## 4. Emerging Directions

### 4.1. AI-Enhanced Tumor Board Support

Medical-focused AI is gaining momentum as a decision-support tool for multidisciplinary tumor boards, where radiology provides a central foundation for oncologic care planning. By synthesizing imaging findings with pathology, genomics, and clinical records, AI has the potential to streamline case preparation, reduce administrative workload, and strengthen collaboration across specialties.

A recent study assessed 102 complex oncology cases, providing ChatGPT-4o with detailed summaries that included imaging, pathology, and prior treatment history [155]. While the model produced plausible recommendations, concordance with tumor board decisions remained modest, highlighting both the promise and the current limitations of LLM-based tumor board support. Similarly, one study tested GPT-3.5 Turbo on the data from 52 patients with non-small cell lung cancer. Each case summary included patient history, imaging reports, and guideline context. The LLM was tasked with recommending one management strategy. The agreement with the decisions of the tumor board reached 76% and the model showed a particularly high consistency in the surgical recommendations [156].

These studies demonstrate the early value of LLMs in handling clinical narratives, but they also reveal an important gap: neither incorporated radiology or pathology images together with text. So far, no published work has combined both imaging and text data in the tumor board setting. Advancing toward true multimodal integration of imaging and clinical data represents a key opportunity for medical AI. Future systems that can simultaneously interpret radiology, pathology, genomic, and clinical information will better capture the multidisciplinary character of tumor boards and provide more effective support for oncology decision making.

### 4.2. AI for Clinical Trial Matching

Clinical trial enrollment depends on accurately identifying patients who meet, often complex and cumbersome, eligibility rules, but this process is time-consuming and is routinely done manually. AI models are now being developed to support this task by analyzing patient records along with the clinical trial enrollment criteria. A recent study showed that LLMs can effectively match patients to clinical trials by interpreting free-text eligibility criteria and comparing them with the patient's EHR data, achieving better performance than traditional rule-based methods [157].

A promising direction for future work is to extend such approaches beyond textual records to include multimodal data sources such as diagnostic images and radiology reports. Since many eligibility criteria depend on imaging findings (e.g., tumor size, metastasis status, disease progression), VLMs that jointly reason over text and images could enable more comprehensive trial matching. Including imaging is important because reports may omit quantitative details, such as exact lesion measurements, vary by reporting style, or overlook subtle findings that are relevant to eligibility. Direct access to images allows AI systems to extract standardized biomarkers and verify reported

impressions, resulting in more accurate and robust clinical trial - patient matching. Integrating these multimodal AI systems into clinical workflows would enable a major advancement in clinical trials.

#### 4.3. AI-Driven Radiology Report Quality Assurance & Error Detection

Quality assurance and error detection are essential components in radiology workflows to maintain the accuracy of the radiology reports and protect patient safety. However, these tasks are both time-consuming and susceptible to human variability. LLMs are increasingly being applied to support this process by automatically detecting inconsistencies and potential errors in imaging findings. For example, a recent large-scale validation study demonstrated that GPT-4 could proofread head CT reports with high sensitivity to both factual and interpretive errors, achieving near-radiologist performance while greatly reducing review time [158].

More recently, multimodal approaches have been explored to cross-check reports directly against the underlying images [159]. A study evaluated VLMs for radiology report error detection by introducing synthetic changes, such as inserted or omitted findings, and found that models with access to chest radiographs and reports outperformed text-only systems. These findings suggest that introducing VLMs in radiology reporting workflows could move quality assurance from retrospective error detection to proactive, image-based validation that strengthens both accuracy and patient safety.

#### 4.4. AI-Driven Complexity Indexing for Radiological Imaging

The U.S. radiology reimbursement system is built around CPT codes and RVUs, but these tools do not reflect the wide variation in interpretive difficulty that often exists among studies assigned the same code [160,161]. In practice, radiologists receive the same payment whether a case is straightforward or requires substantial cognitive effort [160]. Work that demands close comparison with multiple prior studies, careful reconstruction of treatment histories, or nuanced assessment of treatment response typically takes far longer and adds to the growing problem of professional fatigue and burnout [162, 163]. With imaging volumes continuing to rise and staffing pressures becoming more acute, the absence of a way to represent case complexity has created a persistent mismatch between the work performed and the reimbursement assigned.

Recent advances in multimodal AI offer a way to address this gap by generating transparent, quantitative measures of interpretive complexity. These models can integrate patient comorbidities, the number and timing of prior examinations, treatment timelines, the specificity of the clinical question, and image-derived characteristics such as lesion burden, expected treatment effects, and post-operative changes. These elements can then be combined into a single complexity score that reflects the level of interpretive effort likely required. As decision-support tools become more integrated into clinical practice, a validated complexity index has the potential to give departments a clearer picture of workload, provide payers with a more appropriate basis for recognizing high-effort interpretations, and support more consistent, high-quality imaging assessment.

## 5. Conclusions

AI in radiology is steadily moving from experimental promise to practical application. Our experience at Moffitt Cancer Center shows that meaningful progress requires more than technical performance: preferably local deployment, multimodal reasoning, data security, and human oversight are equally critical for building trust and ensuring clinical value. Emerging applications including multimodal AI for tumor board support, clinical trial-patient matching, automated detection of inconsistencies or errors in radiology reports, and the development of imaging complexity indices illustrate how AI can augment decision-making in ways that directly influence patient care. Future success will depend on transparent development, cross-disciplinary collaboration, and alignment with real-world radiology and oncology workflows, ensuring that AI strengthens both diagnostic quality and patient safety.

**Author Contributions:** Conceptualization, I.H., C.L., C.A., and G.R.; writing—original draft preparation, I.H., N.K., S.A., and G.R.; writing—review and editing, I.H., C.L., C.A., N.K., S.A., N.P., M.S., A.Q., R.G., and G.R.; funding acquisition, G.R. and R.G. All authors have read and agreed to the published version of the manuscript.

**Funding:** This study was partly funded by the Department of Diagnostic Imaging and Interventional Radiology, Moffitt Cancer Center, Florida Biomedical Research Award (21B12), NIH/NCI award (U01-CA200464), and NSF Awards 2234836 and 2234468.

**Data Availability Statement:** No new data were created or analyzed in this study. Data sharing is not applicable to this article.

**Acknowledgments:** The authors gratefully acknowledge the Data Services Core at the Moffitt Cancer Center & Research Institute, an NCI-designated Comprehensive Cancer Center (P30-CA076292), for its collaborative support. We further acknowledge our ongoing collaboration with IBIS, Inc., which is funded by three NIH/NCI SBIR awards [75N91021C00042, 75N91023C00041 and 75N91024C00069], addressing the de-identification of PHI and PII, and advancing multimodal machine learning research.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## Abbreviations

The following abbreviations are used in this manuscript:

AI	Artificial Intelligence
NLP	Natural Language Processing
CNN	Convolutional Neural Network
ViT	Visual Transformer
LLM	Large Language Model
VLM	Vision Language Model
PHI	Protected Health Information
EHR	Electronic Health Record
VQA	Visual Question Answering
FL	Federated Learning
HITL	Human-in-the-loop

## References

1. Siewert, B.; Bruno, M.A.; Bourland, J.D.; Slanetz, P.J.; Guillerman, P.; Schwartz, E.S.; Paltiel, H.J.; Hubball, R.; Brook, O.R.; Scanlon, M.H.; et al. Seven Challenges in Radiology Practice: From Declining Reimbursement to Inadequate Labor Force: Summary of the 2023 ACR Intersociety Meeting. *Journal of the American College of Radiology* **2025**, *22*, 129–138. <https://doi.org/10.1016/j.jacr.2024.08.030>.
2. Korfiatis, P.; Kline, T.L.; Meyer, H.M.; Khalid, S.; Leiner, T.; Loufek, B.T.; Blezek, D.; Vidal, D.E.; Hartman, R.P.; Joppa, L.J.; et al. Implementing Artificial Intelligence Algorithms in the Radiology Workflow: Challenges and Opportunities. *Mayo Clinic Proceedings: Digital Health* **2025**, *3*. <https://doi.org/10.1016/j.mcpdig.2024.100188>.
3. Adams, L.C.; Truhn, D.; Busch, F.; Kader, A.; Niehues, S.M.; Makowski, M.R.; Bressemer, K.K. Leveraging GPT-4 for Post Hoc Transformation of Free-text Radiology Reports into Structured Reporting: A Multilingual Feasibility Study. *Radiology* **2023**, *307*. <https://doi.org/10.1148/radiol.230725>.
4. Guellec, B.L.; Lefèvre, A.; Geay, C.; Shorten, L.; Bruge, C.; Haccin-Bey, L.; Amouyel, P.; Pruvo, J.P.; Kuchcinski, G.; Hamroun, A. Performance of an Open-Source Large Language Model in Extracting Information from Free-Text Radiology Reports. *Radiology: Artificial Intelligence* **2024**, *6*. <https://doi.org/10.1148/ryai.230364>.
5. Ranjit, M.; Ganapathy, G.; Srivastav, S.; Ganu, T.; Oruganti, S. RAD-PHI2: Instruction Tuning PHI-2 for Radiology, 2024, [arXiv:cs.CL/2403.09725].
6. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. *Medical Image Computing and Computer-Assisted Intervention – MICCAI* **2015**. [https://doi.org/10.1007/978-3-319-24574-4\\_28](https://doi.org/10.1007/978-3-319-24574-4_28).

7. Ragab, M.G.; Abdulkadir, S.J.; Muneer, A.; Alqushaibi, A.; Sumiea, E.H.; Qureshi, R.; Al-Selwi, S.M.; Alhussian, H. A Comprehensive Systematic Review of YOLO for Medical Object Detection (2018 to 2023). *IEEE Access* **2024**, *12*, 57815–57836. <https://doi.org/10.1109/ACCESS.2024.3386826>.
8. Chen, C.; Isa, N.A.M.; Liu, X. A review of convolutional neural network based methods for medical image classification. *Computers in Biology and Medicine* **2025**, *185*. <https://doi.org/10.1016/j.compbiomed.2024.109507>.
9. Tanno, R.; Barrett, D.G.T.; Sellergren, A.; Ghaisas, S.; Dathathri, S.; See, A.; Welbl, J.; Lau, C.; Tu, T.; Azizi, S.; et al. Collaboration between clinicians and vision–language models in radiology report generation. *Nature Medicine* **2025**, *31*, 599–608. <https://doi.org/10.1038/s41591-024-03302-1>.
10. Deperrois, N.; Matsuo, H.; Ruipérez-Campillo, S.; Vandenhirtz, M.; Laguna, S.; Ryser, A.; Fujimoto, K.; Nishio, M.; Sutter, T.M.; Vogt, J.E.; et al. RadVLM: A Multitask Conversational Vision-Language Model for Radiology, 2025, [arXiv:cs.CV/2502.03333].
11. Xin, Y.; Ates, G.C.; Gong, K.; Shao, W. Med3DVLM: An Efficient Vision-Language Model for 3D Medical Image Analysis, 2025, [arXiv:cs.CV/2503.20047].
12. Gu, D.; Gao, Y.; Zhou, Y.; Zhou, M.; Metaxas, D. RadAlign: Advancing Radiology Report Generation with Vision-Language Concept Alignment, 2025, [arXiv:cs.CV/2501.07525].
13. Park, J.; Kim, S.; Yoon, B.; Choi, K. RadZero: Similarity-Based Cross-Attention for Explainable Vision-Language Alignment in Radiology with Zero-Shot Multi-Task Capability, 2025, [arXiv:cs.CV/2504.07416].
14. Hartsock, I.; Rasool, G. Vision-language models for medical report generation and visual question answering: a review. *Frontiers in Artificial Intelligence* **2024**, *7*. <https://doi.org/doi.org/10.3389/frai.2024.1430984>.
15. Yang, L.; Xu, S.; Sellergren, A.; Kohlberger, T.; Zhou, Y.; Ktena, I.; Kiraly, A.; Ahmed, F.; Hormozdiari, F.; Jaroensri, T.; et al. Advancing Multimodal Medical Capabilities of Gemini, 2024, [arXiv:cs.CV/2405.03162].
16. Wu, C.; Zhang, X.; Zhang, Y.; Wang, Y.; Xie, W. Towards Generalist Foundation Model for Radiology by Leveraging Web-scale 2D&3D Medical Data, 2023, [arXiv:cs.CV/2308.02463].
17. Dancette, C.; Khlaut, J.; Saporta, A.; Philippe, H.; Ferreres, E.; Callard, B.; Danielou, T.; Alberge, L.; Machado, L.; Tordjman, D.; et al. Curia: A Multi-Modal Foundation Model for Radiology, 2025, [arXiv:cs.CV/2509.06830].
18. Machado, L.; Alberge, L.; Philippe, H.; Ferreres, E.; Khlaut, J.; Dupuis, J.; Le Floch, K.; Gatenyo, D.H.; Roux, P.; Grégory, J.; et al. A promptable CT foundation model for solid tumor evaluation. *npj Precision Oncology* **2025**, *9*, 121. <https://doi.org/10.1038/s41698-025-00903-y>.
19. Agrawal, K.K.; Liu, L.; Lian, L.; Nercessian, M.; Harguindeguy, N.; Wu, Y.; Mikhael, P.; Lin, G.; Sequist, L.V.; Fintelmann, F.; et al. Pillar-0: A New Frontier for Radiology Foundation Models. *arXiv preprint arXiv:2511.17803* **2025**.
20. Gupta, V.; Erdal, B.; Ramirez, C.; Floca, R.; Genereaux, B.; Bryson, S.; Bridge, C.; Kleesiek, J.; Nensa, F.; Braren, R.; et al. Current State of Community-Driven Radiological AI Deployment in Medical Imaging. *JMIR AI* **2024**, *3*. <https://doi.org/10.2196/55833>.
21. Marey, A.; Arjmand, P.; Alerab, A.D.S.; Eslami, M.J.; Saad, A.M.; Sanchez, N.; Umair, M. Explainability, transparency and black box challenges of AI in radiology: impact on patient care in cardiovascular radiology. *Egyptian Journal of Radiology and Nuclear Medicine volume* **2024**, *55*. <https://doi.org/10.1186/s43055-024-01356-2>.
22. Koh, D.; Papanikolaou, N.; Bick, U.; Illing, R.; Kahn Jr., C.E.; Kalpathi-Cramer, J.; Matos, C.; Martí-Bonmatí, L.; Miles, A.; Mun, S.K.; et al. Artificial intelligence and machine learning in cancer imaging. *Communications Medicine* **2022**, *2*, 133. <https://doi.org/10.1038/s43856-022-00199-0>.
23. Khalifa, M.; Albadawy, M. AI in diagnostic imaging: Revolutionising accuracy and efficiency. *Computer Methods and Programs in Biomedicine Update* **2024**, *5*, 100146. <https://doi.org/10.1016/j.cmpbup.2024.100146>.
24. Pons, E.; Braun, L.M.M.; Hunink, M.G.M.; Kors, J.A. Natural Language Processing in Radiology: A Systematic Review. *Radiology* **2016**, *279*, 329–343. <https://doi.org/10.1148/radiol.16142770>.
25. Tariq, A.; Banerjee, I.; Trivedi, H.; Gichoya, J. Multimodal Artificial Intelligence Models for Radiology. *BJR | Artificial Intelligence* **2025**, *2*, ubae017. <https://doi.org/10.1093/bjr/ubae017>.
26. Sheller, M.; Edwards, B.; Reina, G.; Martin, J.; Pati, S.; Kotrotsou, A.; Milchenko, M.; Xu, W.; Marcus, D.; Colen, R.; et al. Federated learning in medicine: facilitating multi-institutional collaborations without sharing patient data. *Scientific Reports* **2020**, *10*. <https://doi.org/10.1038/s41598-020-69250-1>.
27. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, L.; Polosukhin, I. Attention Is All You Need. In Proceedings of the Advances in Neural Information Processing Systems, 2017, Vol. 30, pp. 5998–6008.

28. Devlin, J.; Chang, M.W.; Lee, K.; Toutanova, K. BERT: Pre-Training of Deep Bidirectional Transformers for Language Understanding. In Proceedings of the Conference of the North American Chapter of the Association for Computational Linguistics, 2019, Vol. 1, pp. 4171–4186. <https://doi.org/10.18653/v1/N19-1423>.
29. D'Antonoli, T.A.; Stanzione, A.; Bluethgen, C.; Vernuccio, F.; Ugga, L.; Klontzas, M.E.; Cuocolo, R.; Cannella, R.; Koçak, B. Large language models in radiology: fundamentals, applications, ethical considerations, risks, and future directions. *Diagnostic and Interventional Radiology* **2024**, *30*, 80 – 90. <https://doi.org/10.4274/dir.2023.232417>.
30. Smit, A.e.a. Natural Language Processing for Radiology: A Narrative Review. *Radiology* **2023**, *306*. <https://doi.org/10.1148/radiol.230769>.
31. Sahoo, P.; Singh, A.K.; Saha, S.; Jain, V.; Mondal, S.; Chadha, A. A Systematic Survey of Prompt Engineering in Large Language Models: Techniques and Applications, 2025, [arXiv:cs.AI/2402.07927].
32. Schwartz, L.H.; Panicek, D.M.; Berk, A.R.; Li, Y.; Hricak, H. Improving communication of diagnostic radiology findings through structured reporting. *Radiology* **2011**, *260*. <https://doi.org/10.1148/radiol.11101913>.
33. Busch, F.; Hoffmann, L.; dos Santos, D.P.; Makowski, M.R.; Saba, L.; Prucker, P.; Hadamitzky, M.; Navab, N.; Kather, J.N.; Truhn, D.; et al. Large language models for structured reporting in radiology: past, present, and future. *European Radiology* **2025**, *35*. <https://doi.org/10.1007/s00330-024-11107-6>.
34. Hartsock, I.; Araujo, C.; Folio, L.; Rasool, G. Improving Radiology Report Conciseness and Structure via Local Large Language Models. *Journal of Imaging Informatics in Medicine* **2025**. <https://doi.org/10.1007/s10278-025-01510-w>.
35. Reichenpfader, D.; Müller, H.; Denecke, K. A scoping review of large language model based approaches for information extraction from radiology reports. *npj Digital Medicine* **2024**, *7*. <https://doi.org/10.1038/s41746-024-01219-0>.
36. Soni, S.; Gudala, M.; Pajouhi, A.; Roberts, K. RadQA: A Question Answering Dataset to Improve Comprehension of Radiology Reports. In Proceedings of the Proceedings of the Thirteenth Language Resources and Evaluation Conference. European Language Resources Association, 2022, pp. 6250–6259.
37. Yang, X.; Chen, A.; PourNejatian, N.; Shin, H.C.; Smith, K.E.; Parisien, C.; Compas, C.; Martin, C.; Costa, A.B.; Flores, M.G.; et al. A large language model for electronic health records. *npj Digital Medicine* **2022**, *5*. <https://doi.org/10.1038/s41746-022-00742-2>.
38. Kuckelman, I.J.; Yi, P.H.; Bui, M.; Onuh, I.; Anderson, J.A.; MD, A.B.R. Assessing AI-Powered Patient Education: A Case Study in Radiology. *Academic Radiology* **2024**, *31*. <https://doi.org/10.1016/j.acra.2023.08.020>.
39. Ayers, J.W.; Poliak, A.; Dredze, M.; Leas, E.C.; Zhu, Z.; Kelley, J.B.; Faix, D.J.; Goodman, A.M.; Longhurst, C.A.; Hogarth, M.; et al. Comparing Physician and Artificial Intelligence Chatbot Responses to Patient Questions Posted to a Public Social Media Forum. *JAMA Internal Medicine* **2023**, *183*. <https://doi.org/10.1001/jamainternmed.2023.1838>.
40. Hussein, R.; Elkhateeb, A.; Soliman, M.; Yousry, O.; Khedr, H.; Elhawary, H. Evaluating the accuracy and reliability of AI chatbots in patient education on cardiovascular imaging. *Egyptian Journal of Radiology and Nuclear Medicine* **2025**, *56*, 45. <https://doi.org/10.1186/s43055-025-01452-x>.
41. Draelos, R.L.; Afreen, S.; Blasko, B.; Brazile, T.L.; Chase, N.; Desai, D.P.; Evert, J.; Gardner, H.L.; Herrmann, L.; House, A.V.; et al. Large language models provide unsafe answers to patient-posed medical questions, 2025, [arXiv:cs.CL/2507.18905].
42. Özgün Çiçek.; Abdulkadir, A.; Lienkamp, S.S.; Brox, T.; Ronneberger, O. 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation, 2016, [arXiv:cs.CV/1606.06650].
43. Huang, X.; Deng, Z.; Li, D.; Yuan, X.; Fu, Y. MISSFormer: An Effective Transformer for 2D Medical Image Segmentation. *IEEE Transactions on Medical Imaging* **2023**, *42*. <https://doi.org/10.1109/TMI.2022.3230943>.
44. Dai, Y.; Gao, Y.; Liu, F. TransMed: Transformers Advance Multi-Modal Medical Image Classification. *Diagnostics* **2021**, *11*. <https://doi.org/10.3390/diagnostics11081384>.
45. Yamashita, R.; Nishio, M.; Do, R.; Togashi, K. Convolutional Neural Networks: an Overview and Application in Radiology. *Insights into Imaging* **2018**, *9*. <https://doi.org/10.1007/s13244-018-0639-9>.
46. Dosovitskiy, A.; Beyer, L.; Kolesnikov, A.; Weissenborn, D.; Zhai, X.; Unterthiner, T.; Dehghani, M.; Minderer, M.; Heigold, G.; Gelly, S.; et al. An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale. In Proceedings of the International Conference on Learning Representations, 2021.

47. Chen, J.; Mei, J.; Li, X.; Lu, Y.; Yu, Q.; Wei, Q.; et al. TransUNet: Rethinking the U-Net architecture design for medical image segmentation through the lens of transformers. *Medical Image Analysis* **2024**, *97*, 103280. <https://doi.org/https://doi.org/10.1016/j.media.2024.103280>.
48. Hatamizadeh, A.; Tang, Y.; Nath, V.; Yang, D.; Myronenko, A.; Landman, B.; Roth, H.R.; Xu, D. UNETR: Transformers for 3D Medical Image Segmentation. In Proceedings of the Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), 2022, pp. 574–584.
49. Litjens, G.; et al. A Survey on Deep Learning in Medical Image Analysis. *Medical Image Analysis* **2017**, *42*, 60–88. <https://doi.org/10.1016/j.media.2017.07.005>.
50. Teng, Z.; Li, L.; Xin, Z.; Xiang, D.; Huang, J.; Zhou, H.; Shi, F.; Zhu, W.; Cai, J.; Peng, T.; et al. A literature review of artificial intelligence (AI) for medical image segmentation: from AI and explainable AI to trustworthy AI. *Quantitative Imaging in Medicine and Surgery* **14**. <https://doi.org/10.21037/qims-24-723>.
51. Abedalla, A.; Abdullah, M.; Al-Ayyoub, M.; Benkhelifa, E. Chest X-ray pneumothorax segmentation using U-Net with EfficientNet and ResNet architectures. *PeerJ Computer Science* **2021**, *7*, e607. <https://doi.org/10.7717/peerj-cs.607>.
52. Meshram, N.H.; Mitchell, C.C.; Wilbrand, S.; Dempsey, R.J.; Varghese, T. Deep Learning for Carotid Plaque Segmentation using a Dilated U-Net Architecture. *Ultrasonic Imaging* **2020**, *42*, 221–230. <https://doi.org/10.1177/0161734620951216>.
53. Chiu, T.W.; Tsai, Y.L.; Su, S.F. Automatic detect lung node with deep learning in segmentation and imbalance data labeling. *Scientific Reports* **2021**, *11*, 11174. <https://doi.org/10.1038/s41598-021-90599-4>.
54. Zunair, H.; Hamza, A.B. Sharp U-Net: Depthwise convolutional network for biomedical image segmentation. *Computers in Biology and Medicine* **2021**, *136*. <https://doi.org/10.1016/j.compbiomed.2021.104699>.
55. Hirsch, L.; Huang, Y.; Luo, S.; Saccarelli, C.R.; Gullo, R.L.; Naranjo, I.D.; Bitencourt, A.G.V.; Onishi, N.; Ko, E.S.; Leithner, D.; et al. Radiologist-Level Performance by Using Deep Learning for Segmentation of Breast Cancers on MRI Scans. *Radiology: Artificial Intelligence* **2021**, *4*. <https://doi.org/10.1148/ryai.200231>.
56. Azad, R.; Heidari, M.; Shariatnia, M.; Aghdam, E.K.; Karimijafarbigloo, S.; Adeli, E.; Merhof, D. Trans-DeepLab: Convolution-Free Transformer-based DeepLab v3+ for Medical Image Segmentation, 2022, [[arXiv:eess.IV/2208.00713](https://arxiv.org/abs/2208.00713)].
57. Li, S.; Sui, X.; Luo, X.; Xu, X.; Yong, L.; Goh, R.S.M. Medical Image Segmentation using Squeeze-and-Expansion Transformers. In Proceedings of the The 30th International Joint Conference on Artificial Intelligence (IJCAI), 2021.
58. Heidari, M.; Kazerouni, A.; Soltany, M.; Azad, R.; Aghdam, E.K.; Cohen-Adad, J.; Merhof, D. HiFormer: Hierarchical Multi-scale Representations Using Transformers for Medical Image Segmentation. In Proceedings of the 2023 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), 2023, pp. 6191–6201. <https://doi.org/10.1109/WACV56688.2023.00614>.
59. Cheng, P.M.; Montagnon, E.; Yamashita, R.; Pan, I.; Cadrin-Chênevert, A.; Romero, F.P.; Chartrand, G.; Kadoury, S.; Tang, A. Deep Learning: An Update for Radiologists. *RadioGraphics* **2021**, *41*. <https://doi.org/10.1148/rg.2021200210>.
60. Tian, J.; Lee, S.; Kang, K. Faster R-CNN in Healthcare and Disease Detection: A Comprehensive Review. In Proceedings of the 2025 International Conference on Electronics, Information, and Communication (ICEIC), 2025, pp. 1–6. <https://doi.org/10.1109/ICEIC64972.2025.10879615>.
61. Lin, T.Y.; Goyal, P.; Girshick, R.; He, K.; Dollár, P. Focal Loss for Dense Object Detection, 2018, [[arXiv:cs.CV/1708.02002](https://arxiv.org/abs/1708.02002)].
62. Gu, Y.; Lu, X.; Yang, L.; Zhang, B.; Yu, D.; Zhao, Y.; Gao, L.; Wu, L.; Zhou, T. Automatic lung nodule detection using a 3D deep convolutional neural network combined with a multi-scale prediction strategy in chest CTs. *Computers in Biology and Medicine* **2018**, *103*, 220–231. <https://doi.org/10.1016/j.compbiomed.2018.10.011>.
63. Azad, R.; Heidari, M.; Shariatnia, M.; Aghdam, E.K.; Karimijafarbigloo, S.; Adeli, E.; Merhof, D. Advances in medical image analysis with vision Transformers: A comprehensive review. *Medical Image Analysis* **2024**, *91*. <https://doi.org/10.1016/j.media.2023.103000>.
64. Rajpurkar, P.; Irvin, J.; Zhu, K.; Yang, B.; Mehta, H.; Duan, T.; Ding, D.; Bagul, A.; Langlotz, C.; Shpanskaya, K.; et al. CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning, 2017, [[arXiv:cs.CV/1711.05225](https://arxiv.org/abs/1711.05225)].
65. Gheflati, B.; Rivaz, H. Vision Transformers for Classification of Breast Ultrasound Images. In Proceedings of the 2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), 2022, pp. 480–483. <https://doi.org/10.1109/EMBC48229.2022.9871809>.

66. Gao, X.; Khan, M.H.M.; Hui, R.; Tian, Z.; Qian, Y.; Gao, A.; Baichoo, S. COVID-VIT: Classification of Covid-19 from 3D CT chest images based on vision transformer model. In Proceedings of the 2022 3rd International Conference on Next Generation Computing Applications (NextComp), 2022, pp. 1–4. <https://doi.org/10.1109/NextComp55567.2022.9932246>.
67. Tanzi, L.; Audisio, A.; Cirrincione, G.; Aprato, A.; Vezzetti, E. Vision Transformer for femur fracture classification. *Injury* **2022**, *53*, 2625–2634. <https://doi.org/10.1016/j.injury.2022.04.013>.
68. Tajbakhsh, N.; Shin, J.Y.; Gurudu, S.R.; Hurst, R.T.; Kendall, C.B.; Gotway, M.B.; Liang, J. Convolutional Neural Networks for Medical Image Analysis: Full Training or Fine Tuning? *IEEE Transactions on Medical Imaging* **2016**, *35*, 1299–1312. <https://doi.org/10.1109/TMI.2016.2535302>.
69. Moor, M.; Huang, Q.; Wu, S.; Yasunaga, M.; Zakka, C.; Dalmia, Y.; Reis, E.P.; Rajpurkar, P.; Leskovec, J. Med-Flamingo: A Multimodal Medical Few-Shot Learner. *arXiv Preprint arXiv:2307.15189* **2023**.
70. Pellegrini, C.; Özsoy, E.; Busam, B.; Navab, N.; Keicher, M. RaDialog: A Large Vision-Language Model for Radiology Report Generation and Conversational Assistance, 2023.
71. Li, P.; Liu, G.; He, J.; Zhao, Z.; Zhong, S. Masked Vision and Language Pre-Training with Unimodal and Multimodal Contrastive Losses for Medical Visual Question Answering. In Proceedings of the Medical Image Computing and Computer Assisted Intervention (MICCAI), 2023, pp. 374–383. [https://doi.org/10.1007/978-3-031-43907-0\\_36](https://doi.org/10.1007/978-3-031-43907-0_36).
72. Johnson, A.E.; Pollard, T.J.; Berkowitz, S.J.; Greenbaum, N.R.; Lungren, M.P.; Deng, C.y.; Mark, R.G.; Horng, S. MIMIC-CXR, a De-Identified Publicly Available Database of Chest Radiographs with Free-Text Reports. *Scientific Data* **2019**, *6*. <https://doi.org/10.1038/s41597-019-0322-0>.
73. Thawkar, O.; Shaker, A.; Mullappilly, S.S.; Cholakkal, H.; Anwer, R.M.; Khan, S.; Laaksonen, J.; Khan, F.S. XrayGPT: Chest Radiographs Summarization using Medical Vision-Language Models, 2023.
74. Irvin, J.; Rajpurkar, P.; Ko, M.; Yu, Y.; Ciurea-Ilcus, S.; Chute, C.; Marklund, H.; Haghighi, B.; Ball, R.; Shpanskaya, K.; et al. CheXpert: A Large Chest Radiograph Dataset with Uncertainty Labels and Expert Comparison, 2019, [arXiv:cs.CV/1901.07031].
75. Vendrowa, E.; Schonfeldb, E. Understanding transfer learning for chest radiograph clinical report generation with modified transformer architectures. *Heliyon* **2023**, *9*. <https://doi.org/10.1016/j.heliyon.2023.e17968>.
76. Ranjit, M.; Ganapathy, G.; Manuel, R.; Ganu, T. Retrieval Augmented Chest X-Ray Report Generation using OpenAI GPT models, 2023.
77. Lau, J.J.; Gayen, S.; Ben Abacha, A.; Demner-Fushman, D. A Dataset of Clinically Generated Visual Questions and Answers about Radiology Images. *Scientific data* **2018**, *5*, 180251. <https://doi.org/10.1038/sdata.2018.251>.
78. Liu, B.; Zhan, L.M.; Xu, L.; Ma, L.; Yang, Y.F.; Wu, X.M. Slake: A Semantically-Labeled Knowledge-Enhanced Dataset for Medical Visual Question Answering. *IEEE 18th International Symposium on Biomedical Imaging (ISBI)* **2021**, pp. 1650–1654. <https://doi.org/10.1109/ISBI48211.2021.9434010>.
79. Yuan, Z.; Jin, Q.; Tan, C.; Zhao, Z.; Yuan, H.; Huang, F.; Huang, S. RAMM: Retrieval-augmented Biomedical Visual Question Answering with Multi-modal Pre-training, 2023.
80. Zhong, Z.; Wang, Y.; Wu, J.; Hsu, W.C.; Somasundaram, V.; Bi, L.; Kulkarni, S.; Ma, Z.; Collins, S.; Baird, G.; et al. Vision-language model for report generation and outcome prediction in CT pulmonary angiogram. *npj Digital Medicine* **2025**, *8*. <https://doi.org/10.1038/s41746-025-01807-8>.
81. Blankemeier, L.; Cohen, J.P.; Kumar, A.; Van Veen, D.; Gardezi, S.J.S.; Paschali, M.; Chen, Z.; Delbrouck, J.B.; Reis, E.; Truys, C.; et al. Merlin: A vision language foundation model for 3d computed tomography. *Research Square* **2024**, pp. rs–3. <https://doi.org/10.21203/rs.3.rs-4546309/v1>.
82. Ankolekar, A.; Boie, S.; Abdollahyan, M.; Gadaleta, E.; Hasheminasab, S.A.; Yang, G.; Beauville, C.; Dikaios, N.; Kastis, G.A.; Bussmann, M.; et al. Advancing breast, lung and prostate cancer research with federated learning. A systematic review. *npj Digital Medicine* **2025**, *8*. <https://doi.org/10.1038/s41746-025-01591-5>.
83. Health Insurance Portability and Accountability Act of 1996. Pub. L. No. 104-191, 110 Stat. 1936, 1996. Available from: U.S. Government Printing Office, via: <https://www.govinfo.gov/content/pkg/PLAW-104publ191/pdf/PLAW-104publ191.pdf>.
84. General Data Protection Regulation (GDPR). Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016, 2016. Available from: <https://eur-lex.europa.eu/eli/reg/2016/679/oj>.
85. McMahan, B.; Moore, E.; Ramage, D.; Hampson, S.; Arcas, y.B.A. Communication-efficient learning of deep networks from decentralized data. In Proceedings of the Proceedings of Artificial Intelligence and Statistics, 2017, pp. 1273–1282.

86. Darzidehkalani, E.; Ghasemi-rad, M.; van Ooijen, P.V. Federated Learning in Medical Imaging: Part II: Methods, Challenges, and Considerations. *Journal of the American College of Radiology* **2022**, *19*, P755–765. <https://doi.org/10.1016/j.jacr.2022.03.016>.
87. Pati, S.; Baid, U.; Zenk, M.; Edwards, B.; Sheller, M.; Reina, G.A.; Foley, P.; Gruzdev, A.; Martin, J.; Albarqouni, S.; et al. The Federated Tumor Segmentation (FeTS) Challenge, 2021, [arXiv:eess.IV/2105.05874].
88. Che, L.; Zhang, W.; Zhou, M.; Wang, Y.; Ma, H. Multimodal Federated Learning: A Survey. *Sensors* **2023**, *23*, 6986. <https://doi.org/10.3390/s23156986>.
89. Zhang, K.; Song, X.; Zhang, C.; Yu, C. Challenges and future directions of secure federated learning: a survey. *Frontiers in Computer Science* **2022**, *16*, 165817.
90. Wu, R.; Chen, X.; Guo, C.; Weinberger, K.Q. Learning to Invert: Simple Adaptive Attacks for Gradient Inversion in Federated Learning. *arXiv* **2022**. <https://doi.org/10.48550/arXiv.2210.10880>.
91. Koutsoubis, N.; Waqas, A.; Yilmaz, Y.; Ramachandran, R.P.; Schabath, M.B.; Rasool, G. Privacy-preserving Federated Learning and Uncertainty Quantification in Medical Imaging. *Radiology: Artificial Intelligence* **2025**. <https://doi.org/10.1148/ryai.240637>.
92. Adnan, M.; Kalra, S.; Cresswell, J.C.; Taylor, G.W.; Tizhoosh, H. Federated learning and differential privacy for medical image analysis. *Nature* **2021**. <https://doi.org/10.1038/s41598-022-05539-7>.
93. Wei, K.; Li, J.; Ding, M.; Ma, C.; Yang, H.H.; Farokhi, F.; Jin, S.; Quek, T.Q.S.; Vincent Poor, H. Federated Learning With Differential Privacy: Algorithms and Performance Analysis. *IEEE Transactions on Information Forensics and Security* **2020**, *15*, 3454–3469. <https://doi.org/10.1109/TIFS.2020.2988575>.
94. Acar, A.; Aksu, H.; Uluagac, A.S.; Conti, M. A Survey on Homomorphic Encryption Schemes: Theory and Implementation, 2017, [arXiv:cs.CR/1704.03578].
95. Stripelis, D.; Saleem, H.; Ghai, T.; Dhinagar, N.J.; Gupta, U.; Anastasiou, C.; Steeg, G.V.; Ravi, S.; Naveed, M.; Thompson, P.M.; et al. Secure neuroimaging analysis using federated learning with homomorphic encryption. In Proceedings of the SPIE Medical Imaging, 2021. <https://doi.org/10.1117/12.2606256>.
96. Jain, S.; Wallace, B.C. Attention is not Explanation, 2019, [arXiv:cs.CL/1902.10186].
97. Lundberg, S.M.; Lee, S.I. A Unified Approach to Interpreting Model Predictions. In Proceedings of the Advances in Neural Information Processing Systems; Guyon, I.; Luxburg, U.V.; Bengio, S.; Wallach, H.; Fergus, R.; Vishwanathan, S.; Garnett, R., Eds. Curran Associates, Inc., 2017, Vol. 30.
98. Goldshmidt, R.; Horovicz, M. TokenSHAP: Interpreting Large Language Models with Monte Carlo Shapley Value Estimation, 2024, [arXiv:cs.CL/2407.10114].
99. Wei, J.; Wang, X.; Schuurmans, D.; Bosma, M.; Ichter, B.; Xia, F.; Chi, E.; Le, Q.; Zhou, D. Chain-of-Thought Prompting Elicits Reasoning in Large Language Models, 2023, [arXiv:cs.CL/2201.11903].
100. DeepSeek-AI; Guo, D.; Yang, D.; Zhang, H.; Song, J.; Zhang, R.; Xu, R.; Zhu, Q.; Ma, S.; Wang, P.; et al. DeepSeek-R1: Incentivizing Reasoning Capability in LLMs via Reinforcement Learning, 2025, [arXiv:cs.CL/2501.12948].
101. Springenberg, J.T.; Dosovitskiy, A.; Brox, T.; Riedmiller, M. Striving for Simplicity: The All Convolutional Net, 2015, [arXiv:cs.LG/1412.6806].
102. Zhou, B.; Khosla, A.; Lapedriza, A.; Oliva, A.; Torralba, A. Learning Deep Features for Discriminative Localization. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 2921–2929. <https://doi.org/10.1109/CVPR.2016.319>.
103. Selvaraju, R.R.; 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 2017 IEEE International Conference on Computer Vision (ICCV), 2017, pp. 618–626. <https://doi.org/10.1109/ICCV.2017.74>.
104. Wang, H.; Wang, Z.; Du, M.; Yang, F.; Zhang, Z.; Ding, S.; Mardziel, P.P.; Hu, X. Score-CAM: Score-Weighted Visual Explanations for Convolutional Neural Networks. *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)* **2019**, pp. 111–119.
105. Kim, B.; Wattenberg, M.; Gilmer, J.; Cai, C.J.; Wexler, J.; Viégas, F.B.; Sayres, R. Interpretability Beyond Feature Attribution: Quantitative Testing with Concept Activation Vectors (TCAV). In Proceedings of the International Conference on Machine Learning, 2017.
106. Zhang, Y.; Hong, D.; McClement, D.; Oladosu, O.; Pridham, G.; Slaney, G. Grad-CAM helps interpret the deep learning models trained to classify multiple sclerosis types using clinical brain magnetic resonance imaging. *Journal of Neuroscience Methods* **2021**, *353*. <https://doi.org/10.1016/j.jneumeth.2021.109098>.
107. Alghamdi, A.; Alghamdi, K.; Alqahtani, M.; AboHejji, R. Enhance Breast Cancer Diagnosis Using Deep Learning Models On Mammogram Images in Saudi Arabia. In Proceedings of the 2025 International

- Conference on Innovation in Artificial Intelligence and Internet of Things (AIIT), 2025, pp. 1–8. <https://doi.org/10.1109/AIIT63112.2025.11082844>.
108. Maksudov, B.; Curran, K.; Mileo, A. Towards generating more interpretable counterfactuals via concept vectors: a preliminary study on chest X-rays, 2025, [arXiv:eess.IV/2506.04058].
  109. Aflalo, E.; Du, M.; Tseng, S.Y.; Liu, Y.; Wu, C.; Duan, N.; Lal, V. VL-InterpreT: An Interactive Visualization Tool for Interpreting Vision-Language Transformers, 2022, [arXiv:cs.CV/2203.17247].
  110. Parcalabescu, L.; Frank, A. MM-SHAP: A Performance-agnostic Metric for Measuring Multimodal Contributions in Vision and Language Models & Tasks. In Proceedings of the Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). Association for Computational Linguistics, 2023. <https://doi.org/10.18653/v1/2023.acl-long.223>.
  111. Goldshmidt, R. Attention, Please! PixelSHAP Reveals What Vision-Language Models Actually Focus On, 2025, [arXiv:cs.CV/2503.06670].
  112. editorial team, N.M. For trustworthy AI, keep the human in the loop. *Nature Medicine* **2025**, *31*, 3207. <https://doi.org/10.1038/s41591-025-04033-7>.
  113. Luo, M.; Yousefirizi, F.; Rouzrokh, P.; Jin, W.; Alberts, I.; Gowdy, C.; Bouchareb, Y.; Hamarneh, G.; Klyuzhin, I.; Rahmim, A. Physician-in-the-Loop Active Learning in Radiology Artificial Intelligence Workflows: Opportunities, Challenges, and Future Directions. *American Journal of Roentgenology* **2025**. <https://doi.org/10.2214/AJR.25.33364>.
  114. Kaufmann, T.; Weng, P.; Bengs, V.; Hüllermeier, E. A Survey of Reinforcement Learning from Human Feedback, 2024, [arXiv:cs.LG/2312.14925].
  115. Acosta, J.N.; Dogra, S.; Adithan, S.; Wu, K.; Moritz, M.; Kwak, S.; Rajpurkar, P. The Impact of AI Assistance on Radiology Reporting: A Pilot Study Using Simulated AI Draft Reports, 2024, [arXiv:cs.HC/2412.12042].
  116. Awasthi, A.; Le, N.; Deng, Z.; Wu, C.C.; Nguyen, H.V. Enhancing Radiological Diagnosis: A Collaborative Approach Integrating AI and Human Expertise for Visual Miss Correction, 2024, [arXiv:eess.IV/2406.19686].
  117. Chow, J.; Lee, R.; Wu, H. How Do Radiologists Currently Monitor AI in Radiology and What Challenges Do They Face? An Interview Study and Qualitative Analysis. *Journal of Imaging Informatics in Medicine* **2025**. <https://doi.org/10.1007/s10278-025-01493-8>.
  118. Blezek, D.J.; Olson-Williams, L.; Missert, A.; Korfiatis, P. AI Integration in the Clinical Workflow. *Journal of Digital Imaging* **2021**, *34*, 1435–1446. <https://doi.org/10.1007/s10278-021-00525-3>.
  119. Yu, A.C.; Mohajer, B.; Eng, J. External Validation of Deep Learning Algorithms for Radiologic Diagnosis: A Systematic Review. *Radiology: Artificial Intelligence* **2022**, *4*, e210064. <https://doi.org/10.1148/ryai.210064>.
  120. Topff, L.; Groot Lipman, K.B.W.; the ICOVAL, I.C.f.C.I.A.; Guffens, F.; Wittenberg, R.; Bartels-Rutten, A.; van Veenendaal, G.; Hess, M.; Lamerigts, K.; Wakkie, J.; et al. Is the generalizability of a developed artificial intelligence algorithm for COVID-19 on chest CT sufficient for clinical use? Results from the International Consortium for COVID-19 Imaging AI (ICOVAL). *European Radiology* **2023**, *33*, 4249–4258. <https://doi.org/10.1007/s00330-022-09303-3>.
  121. Wang, X.; Liang, G.; Zhang, Y.; Blanton, H.; Bessinger, Z.; Jacobs, N. Inconsistent Performance of Deep Learning Models on Mammogram Classification. *Journal of the American College of Radiology* **2020**, *17*, 796–803. <https://doi.org/10.1016/j.jacr.2020.01.006>.
  122. Zech, J.R.; Badgeley, M.A.; Liu, M.; Costa, A.B.; Titano, J.J.; Oermann, E.K. Variable generalization performance of a deep learning model to detect pneumonia in chest radiographs: A cross-sectional study. *PLOS Medicine* **2018**, *15*, e1002683. <https://doi.org/10.1371/journal.pmed.1002683>.
  123. Larrazabal, A.J.; Nieto, N.; Peterson, V.; Milone, D.H.; Ferrante, E. Gender imbalance in medical imaging datasets produces biased classifiers for computer-aided diagnosis. *Proceedings of the National Academy of Sciences of the United States of America* **2020**, *117*, 12592–12594. <https://doi.org/10.1073/pnas.1919012117>.
  124. Gichoya, J.W.; Banerjee, I.; Bhimireddy, A.R.; Burns, J.L.; Celi, L.A.; Chen, L.; Correa, R.; Dullerud, N.; Ghassemi, M.; Huang, S.; et al. AI recognition of patient race in medical imaging: a modelling study. *The Lancet Digital Health* **2022**, *4*, e406–e414. [https://doi.org/10.1016/S2589-7500\(22\)00063-2](https://doi.org/10.1016/S2589-7500(22)00063-2).
  125. Linardos, A.; Kushibar, K.; Walsh, S.; Gkontra, P.; Lekadir, K. Federated learning for multi-center imaging diagnostics: a simulation study in cardiovascular disease. *Scientific Reports* **2022**, *12*, 3551. <https://doi.org/10.1038/s41598-022-07186-4>.
  126. Sarma, K.V.; Harmon, S.; Sanford, T.; Roth, H.R.; Xu, Z.; Tetreault, J.; Xu, D.; Flores, M.G.; Raman, A.G.; Kulkarni, R.; et al. Federated learning improves site performance in multicenter deep learning without data sharing. *Journal of the American Medical Informatics Association* **2021**, *28*, 1259–1264. <https://doi.org/10.1093/jamia/ocaa341>.

127. Guan, H.; Liu, Y.; Yang, E.; Yap, P.; Shen, D.; Liu, M. Multi-site MRI harmonization via attention-guided deep domain adaptation for brain disorder identification. *Medical Image Analysis* **2021**, *71*, 102076. <https://doi.org/10.1016/j.media.2021.102076>.
128. Liu, S.; Yap, P. Learning multi-site harmonization of magnetic resonance images without traveling human phantoms. *Communications Engineering* **2024**, *3*. <https://doi.org/10.1038/s44172-023-00140-w>.
129. Mallardi, G.; Calefato, F.; Lanubile, F.; Logroscino, G.; Tafuri, B. Diffusion Models for Neuroimaging Data Augmentation: Assessing Realism and Clinical Relevance. *Journal of Medical Systems* **2025**, *49*, 161. <https://doi.org/10.1007/s10916-025-02300-1>.
130. Li, Y.; Vasconcelos, N. REPAIR: Removing Representation Bias by Dataset Resampling. In Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019, pp. 9564–9573. <https://doi.org/10.1109/CVPR.2019.00980>.
131. Zhang, B.H.; Lemoine, B.; Mitchell, M. Mitigating Unwanted Biases with Adversarial Learning. In Proceedings of the Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society. Association for Computing Machinery, 2018, p. 335–340. <https://doi.org/10.1145/3278721.3278779>.
132. Gupta, V.; Erdal, B.; Ramirez, C.; Floca\*, R.O.; Genereaux, B.; Bryson, S.; Bridge, C.; Kleesiek, J.; Nensa, F.; Braren, R.; et al. Current State of Community-Driven Radiological AI Deployment in Medical Imaging. *JMIR AI* **2024**, *3*, e55833. <https://doi.org/10.2196/55833>.
133. Eichelberg, M.; Kleber, K.; Kämmerer, M. Cybersecurity in PACS and Medical Imaging: an Overview. *Journal of Digital Imaging* **2020**, *33*, 1527–1542. <https://doi.org/10.1007/s10278-020-00393-3>.
134. Nguyen, X.V.; Petscavage-Thomas, J.M.; Straus, C.M.; Ikuta, I. Cybersecurity in radiology: Cautionary Tales, Proactive Prevention, and What to do When You Get Hacked. *Current Problems in Diagnostic Radiology* **2025**, *54*, 245–250. <https://doi.org/10.1067/j.cpradiol.2024.07.010>.
135. Pianykh, O.S.; Langs, G.; Dewey, M.; Enzmann, D.R.; Herold, C.J.; Schoenberg, S.O.; Brink, J.A. Continuous Learning AI in Radiology: Implementation Principles and Early Applications. *Radiology* **2020**, *297*, 6–14. <https://doi.org/10.1148/radiol.2020200038>.
136. Strohm, L.; Hehakaya, C.; Ranschaert, E.R.; Boon, W.P.C.; Moors, E.H.M. Implementation of artificial intelligence (AI) applications in radiology: hindering and facilitating factors. *European Radiology* **2020**, *30*, 5525–5532. <https://doi.org/10.1007/s00330-020-06946-y>.
137. Blezek, D.J.; Olson-Williams, L.; Missert, A.; Korfiatis, P. AI Integration in the Clinical Workflow. *Journal of Digital Imaging* **2021**, *34*, 1435–1446. <https://doi.org/10.1007/s10278-021-00525-3>.
138. Kim, B.; Romeijn, S.; van Buchem, M.; Rezazade Mehrizi, M.H.; Grootjans, W. A holistic approach to implementing artificial intelligence in radiology. *Insights into Imaging* **2024**, *15*, 22. <https://doi.org/10.1186/s13244-023-01586-4>.
139. Wallis, A.; McCoubrie, P. The radiology report – are we getting the message across? *Clinical Radiology* **2011**, *66*, 1015–1022. <https://doi.org/10.1016/j.crad.2011.05.013>.
140. Jiang, A.Q.; Sablayrolles, A.; Roux, A.; Mensch, A.; Savary, B.; Bamford, C.; Chaplot, D.S.; de las Casas, D.; Hanna, E.B.; Bressand, F.; et al. Mixtral of Experts, 2024, [arXiv:cs.LG/2401.04088].
141. Jiang, A.Q.; Sablayrolles, A.; Mensch, A.; Bamford, C.; Chaplot, D.S.; de las Casas, D.; Bressand, F.; Lengyel, G.; Lample, G.; Saulnier, L.; et al. Mistral 7B, 2023, [arXiv:cs.CL/2310.06825].
142. Grattafiori, A.; Dubey, A.; Jauhri, A.; Pandey, A.; Kadian, A.; Al-Dahle, A.; Letman, A.; Mathur, A.; Schelten, A.; Vaughan, A.; et al. The Llama 3 Herd of Models, 2024, [arXiv:cs.AI/2407.21783].
143. Team, Q. QwQ-32B: Embracing the Power of Reinforcement Learning, 2025.
144. Park, M.A.; Whelan, C.J.; Ahmed, S.; Boeringer, T.; Brown, J.; Carson, T.L.; Crowder, S.L.; Gage, K.; Gregg, C.; Jeong, D.K.; et al. Defining and addressing research priorities in cancer cachexia through transdisciplinary collaboration. *Cancers* **2024**, *16*, 2364.
145. Han, J.; Harrison, L.; Patzelt, L.; Wu, M.; Junker, D.; Herzig, S.; Berriel Diaz, M.; Karampinos, D.C. Imaging modalities for diagnosis and monitoring of cancer cachexia. *EJNMMI research* **2021**, *11*, 94.
146. Babic, A.; Rosenthal, M.H.; Sundaresan, T.K.; Khalaf, N.; Lee, V.; Brais, L.K.; Loftus, M.; Caplan, L.; Denning, S.; Gurung, A.; et al. Adipose tissue and skeletal muscle wasting precede clinical diagnosis of pancreatic cancer. *Nature Communications* **2023**, *14*, 4317.
147. Go, S.I.; Park, M.J.; Park, S.; Kang, M.H.; Kim, H.G.; Kang, J.H.; Kim, J.H.; Lee, G.W. Cachexia index as a potential biomarker for cancer cachexia and a prognostic indicator in diffuse large B-cell lymphoma. *Journal of cachexia, sarcopenia and muscle* **2021**, *12*, 2211–2219.

148. Tan, S.; Xu, J.; Wang, J.; Zhang, Z.; Li, S.; Yan, M.; Tang, M.; Liu, H.; Zhuang, Q.; Xi, Q.; et al. Development and validation of a cancer cachexia risk score for digestive tract cancer patients before abdominal surgery. *Journal of cachexia, sarcopenia and muscle* **2023**, *14*, 891–902.
149. Argilés, J.M.; López-Soriano, F.J.; Toledo, M.; Betancourt, A.; Serpe, R.; Busquets, S. The cachexia score (CASCO): a new tool for staging cachectic cancer patients. *Journal of cachexia, sarcopenia and muscle* **2011**, *2*, 87–93.
150. Ahmed, S.; Dera, D.; Hassan, S.U.; Bouaynaya, N.; Rasool, G. Failure detection in deep neural networks for medical imaging. *Frontiers in Medical Technology* **2022**, *4*, 919046.
151. Nowak, S.; Theis, M.; Wichtmann, B.D.; Faron, A.; Froelich, M.F.; Tollens, F.; Geißler, H.L.; Block, W.; Luetkens, J.A.; Attenberger, U.I.; et al. End-to-end automated body composition analyses with integrated quality control for opportunistic assessment of sarcopenia in CT. *European radiology* **2022**, *32*, 3142–3151.
152. Ahmed, S.; Parker, N.; Park, M.; Davis, E.W.; Permeth, J.B.; Schabath, M.B.; Yilmaz, Y.; Rasool, G. Multimodal AI-driven Biomarker for Early Detection of Cancer Cachexia, 2025, [arXiv:eess.IV/2503.06797].
153. Mei, X.; Liu, Z.; Robson, P.M.; Marinelli, B.; Huang, M.; Doshi, A.; Jacobi, A.; Cao, C.; Link, K.E.; Yang, T.; et al. RadImageNet: An Open Radiologic Deep Learning Research Dataset for Effective Transfer Learning. *Radiology: Artificial Intelligence* **2022**, *4*, e210315. <https://doi.org/10.1148/ryai.210315>.
154. Aberle, D.R.; Berg, C.D.; Black, W.C.; Church, T.R.; Fagerstrom, R.M.; Galen, B.; et al. The National Lung Screening Trial: Overview and Study Design. *Radiology* **2011**, *258*, 243–253. <https://doi.org/10.1148/radiol.10091808>.
155. Karabuğa, B.; Karaçin, C.; Büyükkör, M.; Bayram, D.; Aydemir, E.; Kaya, O.B.; Yılmaz, M.E.; Çamöz, E.S.; Ergün, Y. The Role of Artificial Intelligence (ChatGPT-4o) in Supporting Tumor Board Decisions. *Journal of Clinical Medicine* **2025**, *14*, 3535. <https://doi.org/10.3390/jcm14103535>.
156. Zabaleta, J.; Aguinagalde, B.; Lopez, I.; Fernandez-Monge, A.; Lizarbe, J.A.; Mainer, M.; Ferrer-Bonsoms, J.A.; de Assas, M. Utility of Artificial Intelligence for Decision Making in Thoracic Multidisciplinary Tumor Boards. *Journal of Clinical Medicine* **2025**, *14*, 399. <https://doi.org/10.3390/jcm14020399>.
157. Jin, Q.; Wang, Z.; Floudas, C.S.; Chen, F.; Gong, C.; Bracken-Clarke, D.; Xue, E.; Yang, Y.; Sun, J.; Lu, Z. Matching patients to clinical trials with large language models. *Nature Communications* **2024**, *15*, 9074. <https://doi.org/10.1038/s41467-024-53081-z>.
158. Kim, S.; Kim, D.; Shin, H.J.; Lee, S.H.; Kang, Y.; Jeong, S.; Kim, J.; Han, M.; Lee, S.; Kim, J.; et al. Large-Scale Validation of the Feasibility of GPT-4 as a Proofreading Tool for Head CT Reports. *Radiology* **2025**, *314*, e240701. <https://doi.org/10.1148/radiol.240701>.
159. Wu, J.; Kim, Y.; Keller, E.C.; Chow, J.; Levine, A.P.; Pontikos, N.; Ibrahim, Z.; Taylor, P.; Williams, M.C.; Wu, H. Exploring Multimodal Large Language Models for Radiology Report Error-checking, 2024, [arXiv:cs.CL/2312.13103].
160. Chen, M.M.; Hirsch, J.A.; Lee, R.K.; Hughes, D.R.; Nicola, G.N.; Rosenkrantz, A.B. Determining the Patient Complexity of Head CT Examinations: Implications for Proper Valuation of a Critical Imaging Service. *Current Problems in Diagnostic Radiology* **2020**, *49*, 177–181. <https://doi.org/10.1067/j.cpradiol.2019.05.007>.
161. Forsberg, D.; Rosipko, B.; Sunshine, J.L. Radiologists' Variation of Time to Read Across Different Procedure Types. *Journal of Digital Imaging* **2017**, *30*, 86–94. <https://doi.org/10.1007/s10278-016-9911-z>.
162. McDonald, R.J.; Schwartz, K.M.; Eckel, L.J.; Diehn, F.E.; Hunt, C.H.; Bartholmai, B.J.; Erickson, B.J.; Kallmes, D.F. The Effects of Changes in Utilization and Technological Advancements of Cross-Sectional Imaging on Radiologist Workload. *Academic Radiology* **2015**, *22*, 1191–1198. <https://doi.org/10.1016/j.acra.2015.05.007>.
163. Chetlen, A.L.; Chan, T.L.; Ballard, D.H.; Frigini, L.A.; Hildebrand, A.; Kim, S.; Brian, J.M.; Krupinski, E.A.; Ganeshan, D. Addressing Burnout in Radiologists. *Academic Radiology* **2019**, *26*, 526–533. <https://doi.org/10.1016/j.acra.2018.07.001>.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.