Human and Artificial Intelligence in Radiology: Current Status, Evidence, Regulation, and Future Perspectives

Zainab Magomedova[1,2], Ekaterina S. Pershina[1,2], Keivan Daneshvar[3], Gerd Nöldge[3], Frank Mosler[3*]

[1]Department of Radiology of The City Clinical Hospital No. 1 named after NI Pirogov, Moscow, Russian Federation

[2]Department of Cardiology, Functional and Ultrasound Diagnostics, Sechenov First Moscow State Medical University (Sechenov University), Moscow, Russian Federation

[3]Department of Diagnostic, Interventional and Pediatric Radiology, Inselspital Bern, University of Bern, Bern, Switzerland

Swiss Journal of Radiology and Nuclear Medicine - www.sjoranm.com - Rosenweg 3 in CH-6340 Baar, Switzerland

Abstract

Artificial intelligence (AI) has rapidly evolved into a transformative force in radiology, complementing human intelligence across the entire imaging workflow. Current applications range from image acquisition and reconstruction to automated detection, quantification, triage, and clinical decision support. Evidence to date demonstrates that AI systems can match or exceed human performance in narrowly defined tasks, particularly in pattern recognition and workflow optimization. However, robust prospective validation, demonstration of clinical impact, and proof of generalizability across institutions and populations remain limited.

Human intelligence continues to play a central role in contextual interpretation, integration of clinical information, ethical judgment, and responsibility for patient care. Rather than replacing radiologists, AI is increasingly viewed as an augmentative tool that enhances diagnostic accuracy, efficiency, and consistency when appropriately implemented.
Regulatory frameworks are evolving in response to these developments. In Europe, the Medical Device Regulation (MDR) and the forthcoming AI Act introduce stricter requirements for transparency, risk classification, post-market surveillance, and human oversight. Comparable regulatory efforts are underway globally, aiming to balance innovation with patient safety, data protection, and accountability. Nonetheless, regulatory heterogeneity and the dynamic nature of adaptive AI systems pose ongoing challenges.

Looking ahead, the future of radiology will be shaped by closer human–AI collaboration, increased emphasis on explainability, continuous learning systems under regulatory control, and higher-quality clinical evidence. Education and training of radiologists in AI literacy will be essential. Ultimately, the successful integration of artificial intelligence into radiology will depend not only on technological progress, but also on evidence-based implementation, clear regulation, and sustained human expertise.

Keywords: artificial intelligence, transparency, radiology

*Corresponding author: Frank Mosler - received: 20.12.2025 - peer reviewed, accepted and published: 31.01.2026

Contents:

1. Introduction

Artificial intelligence (AI) has rapidly evolved into a transformative force in radiology, complementing human intelligence across the entire imaging workflow. Current applications range from image acquisition and reconstruction to automated detection, quantification, triage, and clinical decision support. Evidence to date demonstrates that AI systems can match or exceed human performance in narrowly defined tasks, particularly in pattern recognition and workflow optimization. However, robust prospective validation, demonstration of clinical impact, and proof of generalizability across institutions and populations remain limited.

Human intelligence continues to play a central role in contextual interpretation, integration of clinical information, ethical judgment, and responsibility for patient care. Rather than replacing radiologists, AI is increasingly viewed as an augmentative tool that enhances diagnostic accuracy, efficiency, and consistency when appropriately implemented.

Regulatory frameworks are evolving in response to these developments. In Europe, the Medical Device Regulation (MDR) and the forthcoming AI Act introduce stricter requirements for transparency, risk classification, post-market surveillance, and human oversight. Comparable regulatory efforts are underway globally, aiming to balance innovation with patient safety, data protection, and accountability. Nonetheless, regulatory heterogeneity and the dynamic nature of adaptive AI systems pose ongoing challenges.

Looking ahead, the future of radiology will be shaped by closer human–AI collaboration, increased emphasis on explainability, continuous learning systems under regulatory control, and higher-quality clinical evidence. Education and training of radiologists in AI literacy will be essential. Ultimately, the successful integration of artificial intelligence into radiology will depend not only on technological progress, but also on evidence-based implementation, clear regulation, and sustained human expertise.

2. Methods and Sources

The present article is based on a curated, critical analysis of recent high-quality literature addressing the role of artificial intelligence (AI) in radiology. The purpose of this chapter is not to provide a systematic review in the strict methodological sense, but rather to transparently describe the sources used and to analyze how contemporary publications conceptualize, evaluate, and contextualize AI within clinical radiology. Emphasis is placed on methodological rigor, evidence generation, human–AI interaction, and regulatory framing.

The selected time frame (2022–2025) captures a period that followed an initial phase of considerable enthusiasm surrounding AI in radiology. During the years preceding this interval, AI was frequently portrayed as a disruptive technology with the potential to fundamentally transform diagnostic imaging, often accompanied by claims of near-human or superhuman performance. More recent publications, however, reflect a noticeable shift toward a more cautious and sober assessment. This phase is characterized by increased attention to real-world performance, unintended consequences of AI deployment, limitations of retrospective evidence, and the growing influence of regulatory requirements (11).

Against this background, the present analysis aims to examine how leading journals and expert groups currently approach AI in radiology, how evidence is generated and reported, and where persistent gaps and inconsistencies remain. The overarching perspective is radiological and clinical, with patient safety, accountability, and feasibility of implementation taking precedence over technological optimism.

2.1 Literature identification and selection strategy

The literature corpus consists of twelve peer-reviewed publications published between 2022 and 2025. Sources were selected from internationally recognized journals with high relevance for clinical radiology, medical imaging research, digital medicine, and health technology assessment. These include Nature Medicine, Radiology, The Lancet Digital Health, European Radiology, npj Digital Medicine, Insights into Imaging, The British Journal of Radiology, and Value in Health (1–12).

Selection criteria focused on publications that met at least one of the following conditions:
(a) presentation of original clinical or reader-based evidence on AI performance in radiology,
(b) methodological or reporting frameworks for clinical evaluation of AI systems,
(c) health-economic evaluation standards applicable to AI-based interventions, or
(d) regulatory and legal analyses with direct relevance to radiological practice.

Purely technical machine-learning papers without clinical validation, as well as non–peer-reviewed industry reports, were deliberately excluded.

The final selection reflects four thematic clusters: clinical evidence and human–AI interaction (7, 8, 10, 12), methodological and reporting standards (4 – 6), regulatory and governance perspectives (7 – 10, 12), and critical commentaries offering a meta-level appraisal of the current state of AI in radiology (11). This approach allows a balanced view across the AI life cycle, from development and evaluation to implementation and oversight.

2.2 Types of publications and study designs

The analyzed literature demonstrates substantial heterogeneity with regard to publication type and study design. Original clinical evidence is predominantly derived from retrospective cohort studies and reader studies, often using enriched or curated datasets. A representative example is the large retrospective screening study evaluating AI as an independent or assisting reader in breast cancer screening, which relies on historical mammography data with long-term follow-up (2). Similarly, real-world validation studies in specific disease contexts, such as multiple sclerosis MRI monitoring, remain largely retrospective and context-specific (3).

Reader studies assessing human–AI interaction frequently employ simulated reading environments or controlled experimental designs. While these approaches allow detailed analysis of performance metrics and behavioral effects, they inherently differ from routine clinical conditions (1). Prospective randomized trials remain rare, and when present, are often limited to narrow use cases or specific screening settings.

A substantial portion of the literature consists of methodological guidance documents and reporting standards, including frameworks for clinical evaluation (4), early-stage decision support assessment (5), and health-economic reporting (6). These publications are normative in nature and aim to raise the methodological bar for future studies rather than to provide empirical performance data.

Regulatory and legal analyses form another important category. These papers interpret evolving regulatory frameworks, particularly within the European context, and translate legal requirements into practical implications for radiologists and healthcare institutions (7, 10, 12). Finally, critical commentaries synthesize existing evidence and explicitly challenge prevailing assumptions about efficiency gains, economic benefits, and the transformative impact of AI in routine radiology (11).

2.3 Conceptualization of artificial intelligence in the literature

Across the analyzed publications, AI is consistently conceptualized as a task-specific tool rather than a general or autonomous diagnostic entity. Most studies focus on narrowly defined applications such as lesion detection, triage, quantification, or second-reader support. This reflects both technical realities and regulatory constraints, as fully autonomous diagnostic systems remain neither legally permissible nor clinically validated.

Definitions of AI vary in specificity, ranging from broad descriptions encompassing machine learning and deep learning to more ex-

plicit distinctions between conventional algorithms, deep neural networks, and, more recently, large language models (LLMs) (9). However, even when LLMs are discussed, they are framed as adjunctive tools for reporting, documentation, or workflow support rather than primary diagnostic decision-makers.

Functionally, AI systems are most commonly positioned as assistive technologies. The second-reader paradigm, particularly in screening contexts, represents a recurring theme and is often cited as a realistic and regulatorily acceptable use case (2, 11). Workflow-oriented applications, such as protocoling, image reconstruction, or administrative automation, are acknowledged as potentially impactful but remain underrepresented in empirical studies.

Importantly, none of the analyzed sources advocate for the removal of the radiologist from the diagnostic process. On the contrary, explicit emphasis is placed on human oversight, contextual interpretation, and accountability, reinforcing the notion of AI as an augmentative rather than substitutive technology.

2.4 Evaluation methodology and evidence standards

Performance evaluation in the reviewed literature relies heavily on conventional diagnostic metrics, including area under the receiver operating characteristic curve, sensitivity, specificity, and recall rates. While these measures are well established, their clinical relevance is often limited when used in isolation. Improvements in such metrics do not necessarily translate into better patient outcomes, reduced morbidity, or meaningful efficiency gains (18, 19).

Comparisons between human readers and AI systems, or between unassisted and AI-assisted radiologists, reveal heterogeneous effects. A large-scale reader study demonstrated substantial inter-individual variability in the impact of AI assistance, with some radiologists experiencing performance improvements and others showing deterioration, particularly in the presence of AI errors (1). These findings challenge the assumption that AI uniformly benefits less experienced readers or consistently improves overall performance.

External validation remains a major weakness. Many studies rely on single-center datasets or vendor-specific systems, limiting generalizability across institutions, populations, and imaging protocols. Dataset shift and hidden stratification are rarely addressed in a systematic manner.

Methodological guidance documents attempt to address these shortcomings. The DECIDE-AI framework provides structured recommendations for early-stage clinical evaluation of AI-based decision support systems, emphasizing transparency, contextual description, and staged evidence generation (5). Similarly, CHEERS-AI extends established health-economic reporting standards to AI interventions, highlighting the need to explicitly account for algorithm behavior, learning effects, and implementation costs (6). Despite their availability, adherence to these frameworks in empirical studies remains inconsistent.

2.5 Human–AI interaction and the role of the radiologist

The analyzed literature consistently assigns a central role to the radiologist within AI-augmented workflows. Radiologists are portrayed as integrators of imaging findings, clinical context, and patient-specific considerations, functions that remain beyond the capabilities of current AI systems.

Human–AI interaction studies reveal both potential benefits and risks. While AI may support decision-making in specific tasks, it can also introduce automation bias, authority bias, and overreliance, particularly when AI outputs are presented without adequate uncertainty information or calibration (1, 11). Empirical evidence indicates that incorrect AI suggestions can negatively influence radiologist performance, increasing error rates compared with unassisted reading (1).

Concerns regarding deskilling are acknowledged but not uniformly supported by evidence. Rather than a loss of expertise, the literature suggests a redistribution of cognitive effort, with radiologists shifting from pure detection tasks toward supervision, validation, and consultation. However, this shift has implications for training, work-load, and professional responsibility that are only partially addressed in current studies.

2.6 Regulatory and governance perspectives

Regulatory considerations occupy a prominent position in recent literature. In the European context, AI systems used in radiology are consistently classified as high-risk medical devices, subject to stringent requirements under the Medical Device Regulation and the forthcoming AI Act (7, 8, 10, 12). Core principles include risk management throughout the life cycle, robust data governance, transparency, and mandatory human oversight.

Several publications emphasize that regulatory compliance is not merely a legal obligation but a determinant of study design and implementation strategy. Adaptive systems, continuous learning, and post-market performance monitoring pose particular challenges, as they conflict with traditional static approval models (7, 9).

Legal analyses further highlight issues of liability and accountability. Radiologists retain ultimate responsibility for diagnostic decisions, even when AI systems are involved, reinforcing the need for clear protocols defining the scope and limits of AI use (12). Similar regulatory trends are observed globally, suggesting increasing convergence toward risk-based oversight rather than permissive innovation.

2.7 Methodological and structural limitations across the literature

Despite notable progress, several structural limitations persist across the analyzed sources. Publication bias toward positive results remains likely, as negative or neutral findings are under-represented. Many studies address narrowly defined tasks that may not reflect the complexity of routine radiological practice.

Economic evidence is particularly limited. Claims of efficiency gains and cost reduction are often speculative and rarely supported by comprehensive economic modeling or prospective evaluation (6, 11). This gap undermines the business case for widespread AI adoption and contributes to the growing sense of disillusionment following earlier hype.

Moreover, the fragmentation of evidence across clinical, technical, and regulatory domains hampers integrated assessment. Few studies simultaneously address performance, workflow impact, economic implications, and legal feasibility.

2.8 Rationale for integrating these sources

The selected literature provides a coherent snapshot of the current state of AI in radiology, characterized by a transition from enthusiasm to methodological realism. Together, these sources illustrate how AI is increasingly framed as a supportive technology whose value depends on rigorous evaluation, thoughtful implementation, and robust governance.

By integrating clinical studies, methodological frameworks, regulatory analyses, and critical commentaries, this chapter establishes the foundation for subsequent discussion. It highlights both the tangible progress achieved and the substantial work that remains necessary to ensure that AI contributes meaningfully and safely to radiological practice.

3. Radiology Today: Clinical Role and Value Contribution

3.1 Radiology as a clinical discipline in a value-driven era

Radiology has long been central to modern healthcare, yet its clinical role has not always been visible in proportion to its impact. Over the last decades, imaging has moved from a supportive diagnostic tool to a cornerstone of patient pathways, influencing diagnosis, staging, treatment planning, monitoring, and increasingly prognostication. At the same time, radiology has been affected by pressures that are now familiar across many health systems: rising demand, workforce shortages, cost containment, and the shift toward value-based healthcare models. These developments have triggered an important reappraisal of what radiology contributes—beyond throughput, report volume, and technical excellence—and how that contribution should be articulated and measured.

Value-based healthcare is often described as maximizing patient outcomes relative to cost.

In this context, radiology's value is not merely determined by the accuracy of image interpretation, but by its measurable influence on clinical decision-making, patient experience, and down-stream outcomes. A multisociety perspective has emphasized that radiology must be viewed as a clinical discipline whose value extends beyond the production of diagnostic reports and includes consultative expertise, stewardship of appropriate imaging, and quality improvement across the care continuum. (14, 15, 22) Similar arguments have been expressed in broader medical discourse, underscoring that imaging is embedded in clinical decision systems and should be evaluated accordingly. (19)

While European radiology has produced several structured frameworks on value and professional identity, the underlying themes are not uniquely European. Rather, they reflect global challenges: maintaining quality and access under increasing demand, demonstrating relevance within multidisciplinary care, and ensuring that radiology remains clinically integrated rather than commoditized. (18, 19) In this chapter, the contemporary clinical role of radiology is analyzed through the lens of value contribution, focusing on three interconnected domains: clinical impact on decision-making, radiologists as consultants and integrators, and patient-centered value including communication and experience. In parallel, the chapter addresses how value can be measured and where limitations in current evidence and practice remain.

3.2 Clinical impact: radiology as a driver of diagnosis and management

Radiology's most direct contribution to value lies in its influence on diagnostic accuracy and clinical management. Imaging findings frequently determine whether a patient is admitted or discharged, treated conservatively or invasively, and whether a malignancy is staged as resectable or metastatic. In acute care, radiology is pivotal in time-critical decisions such as stroke treatment, trauma triage, and suspected pulmonary embolism. In oncology, imaging guides diagnosis, staging, therapy response assessment, and surveillance. These roles are widely accepted in clinical practice; however, the challenge is that radiology's impact is often diffuse and distributed across multiple decisions, making it difficult to quantify using single metrics.

A value-based framing therefore requires moving beyond "test performance" and considering radiology's effect on downstream outcomes. The multisociety perspective in Radiology emphasizes that radiology creates value through appropriate imaging selection, accurate and timely diagnosis, and the reduction of diagnostic uncertainty. (14, 15, 22) In practice, radiology can shorten time to diagnosis, prevent unnecessary procedures, and improve patient stratification. Yet these benefits may not be captured by traditional departmental key performance indicators such as report turn-around time or scanner utilization.

Moreover, radiology is increasingly intertwined with clinical pathways and guidelines. Imaging appropriateness, radiation safety, and protocol optimization represent value contributions that occur before interpretation even begins. These elements are especially relevant in global healthcare settings where access to imaging is unequal, resources are limited, and the balance between benefit and cost is particularly delicate. A mature value narrative must therefore include not only high resource settings with advanced imaging infrastructure, but also low- and middle-income contexts where radiology may have a different role—sometimes more focused on basic access and diagnostic availability.

In multidisciplinary care, radiology's value becomes more visible. Imaging is a shared language between specialties, and radiologists provide interpretive expertise that can prevent miscommunication and ensure that imaging findings are translated into actionable decisions. A summary of the ESR International Forum highlights that radiology's role in multidisciplinary approaches is not optional but fundamental, and that visibility and integration are essential if radiology is to contribute optimally to patient care. (17) While this discussion emerges from a European forum, the same logic applies globally: the radiologist's clinical role is strongest when radiology is embedded in teams rather than operating as an isolated reporting service.

3.3 The radiologist as consultant and integrator: beyond the report

One of the clearest contemporary shifts in radiology is the reemphasis on the radio-

logist as a clinical consultant. This role is sometimes under-appreciated because it is difficult to measure and often not formally reimbursed. Yet consultative work is a major mechanism through which radiologists create value: recommending the most appropriate imaging, advising on protocol selection, clarifying findings for referring clinicians, and contributing to complex decision-making in multidisciplinary conferences.

A recent analysis in the Journal of the American College of Radiology addresses the value of radiology consultation in terms of effort allocation, clinical impact, and *"untapped opportunities."* (14, 15, 22) The authors frame consultation as a meaningful and underutilized component of radiology practice. This is important for value-based radiology because it challenges a narrow view in which radiology's output is reduced to a written report. Consultation can prevent unnecessary imaging, avoid repeated examinations, improve interpretation accuracy through clinical context, and support decision-making where imaging findings are ambiguous or unexpected.
The consultative role also strengthens radiology's identity as a clinical specialty. The ESR white paper on the changing world of healthcare highlights that radiologists must maintain clinical visibility, participate actively in patient pathways, and engage in communication that demonstrates value to both clinicians and patients. (18, 19) Although this white paper is European in origin, its relevance is global: in many health systems, radiology is vulnerable to commoditization when radiologists are perceived as "report producers" rather than clinical experts.

However, the consultative role is not uniformly implemented worldwide. In some regions, radiologists are physically colocated with clinical teams and participate in ward rounds, tumor boards, and clinical conferences. In other settings, radiology is increasingly remote, with outsourcing, teleradiology, and distributed reporting models. These models may improve access and efficiency, but they risk weakening clinical integration if consultation is not explicitly preserved. Value-based radiology therefore requires intentional structures that enable consultation, such as dedicated clinician communication channels, protected time for multidisciplinary meetings, and recognition of consultation as a measurable service.

3.4 Patient-centered value: communication, understanding, and trust

Radiology's value is not limited to clinicians and health systems; it also extends directly to patients. Historically, radiology has often been a "hidden" specialty, with limited patient contact and little visibility in patient experience narratives. Yet patients increasingly access their imaging reports through electronic health records, and expectations of transparency and communication have risen across healthcare. This shift has implications for radiology's role, responsibilities, and potential value contribution.

Direct communication between radiologists and patients has been studied as a mechanism for improving report quality. A study in European Radiology reported that direct communication can improve the quality of imaging reports. (18, 19) While the precise pathways of this effect can be debated—ranging from improved clinical context to increased accountability—the broader implication is that patient-facing radiology is not merely a "soft skill" but may influence diagnostic clarity and relevance. Communication can also reduce anxiety, correct misunderstandings, and strengthen trust in imaging-based decisions.

Patient perceptions of radiology value have also been explored through surveys. The ESR value-based radiology subcommittee reported results from a patient survey addressing how value is perceived in relation to radiology. (14, 15, 22) Such work is important because value-based healthcare is, at least in principle, patient-centered. If radiology is to demonstrate value, it must understand what patients consider valuable: timely access, clear explanations, respectful interaction, safety, and the sense that imaging contributes meaningfully to care rather than being a routine or redundant step.

Importantly, patient-centered value varies internationally. In some health systems, patients may have direct access to radiologists and structured opportunities for consultation; in others, radiologists remain largely invisible. Cultural expectations also differ: some patients prefer detailed explanations, while others may defer to clinicians. A balanced international perspective therefore avoids prescribing a single model and instead recognizes that patient-centered radiology

must be adapted to local norms and infrastructure.

Nonetheless, the general direction is clear: radiology's value proposition is strengthened when radiologists engage with patients as stake-holders. This does not imply that every radiologist must become a front-line communicator in all settings, but it does suggest that radiology departments should develop strategies for patient communication, report clarity, and accessibility.

3.5 Frameworks for value: defining, measuring, and improving radiology's contribution

The concept of "value" risks becoming rhetorical unless it is linked to measurable and actionable frameworks. Radiology has increasingly adopted value-based language, but implementation requires concrete metrics and quality improvement mechanisms. The ESR has provided structured perspectives on value-based radiology, including discussions on what radiology societies are doing and what future directions should be pursued. (14, 15, 22) These frameworks emphasize that value is multidimensional, involving clinical outcomes, safety, patient experience, appropriateness, efficiency, and professional engagement.

A central challenge is that radiology's value is often indirect. For example, an accurate report may prevent unnecessary surgery, but the avoided harm may not be captured as a radiology metric. Similarly, imaging stewardship may reduce unnecessary examinations, but the "success" is the absence of imaging rather than increased volume. This creates tension with traditional productivity metrics that reward throughput rather than appropriateness.

Feedback mechanisms represent one practical route to value improvement. A recent paper on feedback in radiology describes feedback as an essential tool for improving user experience and delivering value-based care. (14, 15, 22) Feedback can be directed toward multiple stakeholders: referring clinicians, radiologists, technologists, and patients. It can address diagnostic accuracy, report clarity, communication, turnaround time, and appropriateness. Importantly, feedback systems can convert abstract value concepts into operational quality improvement processes.

The multisociety perspective on value-based radiology also emphasizes that radiology must demonstrate its impact through evidence and quality measurement rather than relying on assumptions of importance. (14, 15, 22) In practice, this may involve adopting metrics such as:

- appropriateness and guideline adherence,
- clinically actionable report elements,
- discrepancy tracking and learning systems,
- patient satisfaction and understanding,
- participation in multidisciplinary decision-making,
- time-to-treatment or pathway efficiency measures.

From an international standpoint, the choice of metrics should reflect local priorities. In resource-limited settings, value may be measured through improved access and reduced diagnostic delay. In high-resource systems, value may be measured through appropriateness, cost-effectiveness, and avoidance of unnecessary downstream interventions.

3.6 Radiology identity and professional visibility: maintaining relevance in modern healthcare

The identity of radiology as a clinical specialty is closely linked to value contribution. A survey among ESR full radiologist members explored professional identity and role perception, offering insight into how radiologists view their position in healthcare. (20) While such surveys reflect a specific membership population, they highlight broader professional concerns: maintaining clinical relevance, avoiding commoditization, and ensuring that radiologists are recognized as physicians with interpretive and consultative expertise.

The ESR white paper further elaborates on the radiologist's role in a changing healthcare environment, emphasizing that radiologists must remain clinically engaged, participate in decision-making, and adapt to evolving expectations. (13) These perspectives align with global concerns about workforce shortages, increasing imaging demand, and the need for radiology to maintain both quality and accessibility.

Radiology's identity is also shaped by how it is organized. Departmental integration with clinical services, training structures, and institutional culture all influence whether radiologists are visible as clinical partners. In systems where radiology is primarily service-oriented and remote, radiologists may be less involved in direct clinical dialogue. In contrast, in systems with strong multi-disciplinary integration, radiologists may be perceived as indispensable contributors to patient care.

A key point is that radiology's value is not self-evident to all stakeholders. Hospital administrators may focus on cost and throughput, clinicians may focus on availability and report clarity, and patients may focus on understanding and reassurance. Value-based radiology therefore requires active communication of radiology's contributions, supported by evidence and quality improvement.

3.7 Challenges and limitations: evidence gaps, measurement problems, and implementation barriers

While the value narrative is compelling, it must be tempered by realism. Several limitations persist in how radiology value is currently conceptualized and measured.

First, evidence linking radiology interventions to patient outcomes is often indirect. While it is intuitive that accurate imaging improves care, rigorous studies demonstrating downstream outcomes are difficult to conduct. Imaging is embedded within complex clinical pathways, and isolating radiology's independent effect can be methodologically challenging. As a result, many value arguments rely on plausibility and expert consensus rather than definitive outcome trials.

Second, economic evaluation is frequently underdeveloped. Value-based healthcare is inherently tied to cost-effectiveness, yet radiology economics can be complex. Costs are distributed across equipment, staffing, maintenance, and downstream interventions. Moreover, imaging can both increase and decrease costs: it may reduce unnecessary procedures, but it may also detect incidental findings that generate additional testing. A mature value framework must acknowledge these complexities rather than assuming that imaging always reduces cost.

Third, measurement systems may incentivize the wrong behaviors. If radiology departments are evaluated primarily by throughput and turnaround time, radiologists may have limited incentive or time for consultation, multidisciplinary engagement, and patient communication. Yet these are precisely the activities that strengthen radiology's value contribution. Aligning incentives with value therefore requires institutional commitment and structural support.

Fourth, international variability complicates generalization. Health systems differ in reimbursement, referral patterns, imaging access, and professional roles. A strategy that improves value in one system may not translate directly to another. For example, patient-facing radiology communication may be feasible in some contexts but not in high-volume settings with severe work-force shortages. Similarly, consultation models depend on institutional culture and clinical workflow.

Finally, the shift toward value-based radiology may encounter resistance if it is perceived as an administrative burden rather than a clinical opportunity. The success of value-based initiatives depends on radiologists seeing them as tools for improving care and strengthening professional identity, not merely as reporting requirements.

3.8 Radiology's value proposition today

Radiology today is best understood as a clinical discipline that contributes value across the patient pathway. Its impact extends from accurate diagnosis and management guidance to consultation, multidisciplinary integration, patient communication, and stewardship of appropriate imaging. The transition toward value-based health-care provides both a challenge and an opportunity: radiology must demonstrate its contribution in measurable terms, but it can also strengthen its clinical identity by emphasizing roles that go beyond report production.

Internationally, radiology's value contribution is shaped by local healthcare structures, workforce realities, and cultural expectations. Nonetheless, the core elements of value appear consistent: clinical relevance, integration, communication, safety, and outcome-oriented practice. Frame-works and

professional guidance support this shift, but further work is needed to build robust measurement systems, generate outcome-linked evidence, and align incentives with value. (11 – 15)
Ultimately, radiology's future role will depend on its ability to remain clinically visible, evidence-driven, and patient-centered—ensuring that imaging continues to serve not only diagnostic accuracy, but meaningful improvement in patient care.

4. Radiation Protection as Culture (Technology, Behavior, Organization)

Radiation protection in radiology is increasingly recognized not merely as a collection of technical rules, but as a comprehensive culture that integrates technology, human behavior, and organizational structures. This cultural perspective connects the classical principles of radiation protection—justification, optimization (including the ALARA principle), and dose limitation—with routine clinical practice, clinical decision-making, and leadership within radiology departments. (24) Such an approach reflects the understanding that radiation safety is shaped by everyday professional actions and institutional priorities rather than by technology alone.
This perspective is consistent with international recommendations, which emphasize justification, optimization, and dose limitation as the foundational principles of radiation protection practice. The International Commission on Radiological Protection (ICRP) explicitly frames these principles within a system that requires professional responsibility, education, and organizational support to be effective in clinical settings. (13, 25)

4.1 Technology: Optimization, Automation, and Monitoring

From a technological standpoint, radiation protection relies on optimized imaging protocols and effective dose management systems. Advances in computed tomography and other imaging modalities have introduced automatic exposure control, iterative reconstruction algorithms, and protocol standardization strategies aimed at preserving diagnostic image quality while minimizing patient exposure. However, large-scale studies demonstrate substantial variability in radiation doses between institutions, often attributable to protocol selection, parameter settings, and workflow differences rather than inherent equipment limitations. (26) This variability highlights both the potential and the limitations of purely technical dose-reduction strategies.
Dose monitoring systems, supported by digital imaging and hospital information infrastructures, enable systematic collection and analysis of radiation exposure data. These platforms facilitate benchmarking against diagnostic reference levels (DRLs), identification of outliers, and implementation of corrective actions. When combined with education and structured quality improvement processes, dose monitoring and audit-and-feedback mechanisms have been shown to reduce radiation exposure without compromising diagnostic performance. (27) International organizations, including the International Atomic Energy Agency (IAEA), recommend such systematic approaches as essential components of medical radiation protection programs. (28)

4.2 Behavior: Awareness, Training, and Professional Responsibility

Technology alone cannot ensure radiation safety without informed and deliberate human action. Persistent variability in radiation doses for similar CT examinations across institutions underscores the central role of user-dependent decisions, such as protocol selection and parameter adjustment, in determining patient exposure. (26) These findings indicate that suboptimal practices are often driven by gaps in training, awareness, or routine habits rather than by technical constraints.

A robust radiation protection culture therefore requires continuous education and professional development for radiologists, radiologic technologists, and medical physicists. Professional and international organizations emphasize that radiation protection is an ethical obligation and an integral part of high-quality clinical care. Educational initiatives should address radiation risk communication, evidence-based modality selection, and the application of appropriateness criteria. Behavioral interventions, including structured audits, feedback systems, and collaborative quality improvement initiatives, have demonstrated measurable reductions in unnecessary radiation exposure while maintaining clinical effectiveness. (27)

4.3 Organization: Governance, Processes, and Safety Culture

At the organizational level, embedding radiation protection into governance structures and standard operating procedures promotes consistency, accountability, and sustainability. Designated radiation protection officers, medical physics experts, and multidisciplinary committees play a key role in protocol harmonization, incident reporting, and performance monitoring. Such formal structures support a just and learning safety culture, in which staff are encouraged to report near-miss events and quality concerns without fear of punitive consequences.

International safety standards and regulatory frameworks further reinforce the need to integrate radiation protection into broader health care quality management systems rather than treating it as an isolated compliance requirement. The IAEA safety guidance on medical uses of ionizing radiation provides a comprehensive framework for implementing coordinated technical, educational, and managerial measures to protect patients, workers, and the public. (28)

4.4 Synthesis: Radiation Protection as Culture

Radiation protection in radiology represents a complex socio-technical system. Technology supplies the tools for dose optimization and monitoring; professional behavior translates knowledge into daily practice; and organizational structures ensure reliability, accountability, and continuous learning. When these elements are aligned and supported by international standards and evidence-based governance, radiation protection evolves beyond regulatory compliance into a pervasive culture of safety that enhances patient care and professional integrity in radiology.

5. AI in Radiology: Application Areas and Evidence

Artificial intelligence (AI) has progressed in radiology from experimental prototypes to clinically deployed systems across multiple application domains. Current implementations already demonstrate measurable efficiency gains and task-specific performance improvements, while simultaneously highlighting the necessity of contextual evaluation, continuous monitoring, and sustained human oversight (29, 30)

5.1 Image Interpretation and Diagnostic Support

The most mature and widely studied AI applications in radiology focus on image interpretation, particularly in high-volume examinations such as chest radiography, mammography, and CT. Deep learning algorithms have demonstrated diagnostic performance comparable to expert radiologists in specific, well-defined tasks.

In mammography, a large retrospective study by McKinney et al. showed that a deep learning system reduced both false-positive and false-negative rates compared with human readers across datasets from the United States and the United Kingdom (31). Importantly, the study emphasized that AI performance varied across populations and imaging settings, reinforcing the need for local validation before clinical deployment (31).

For chest X-ray interpretation, commercially deployed systems such as those developed by Annalise.ai are based on multi-label deep learning models trained to detect dozens of radiographic findings simultaneously. Clinical validation studies have demonstrated improved sensitivity for certain pathologies when AI is used as a second reader, particularly in emergency and high-throughput settings (32, 33). However, these gains are task-specific and depend strongly on prevalence, case mix, and reader experience.

5.2 Workflow Automation and Reporting Efficiency

Beyond diagnosis, AI has shown significant impact in workflow automation and reporting efficiency. Natural language processing (NLP) and generative AI techniques are increasingly used for report structuring, auto-completion, and clinical summarization. Studies conducted in academic radiology departments demonstrate that AI-assisted reporting can reduce reporting times while maintaining diagnostic accuracy, particularly for standardized examinations such as trauma CT or chest imaging (34).

Nevertheless, evidence also indicates that unchecked automation may introduce new risks, including propagation of template

errors and reduced critical reflection. Consequently, professional societies emphasize that AI-generated text must remain assistive rather than autonomous, with final responsibility residing unequivocally with the radiologist (29).

5.3 Radiotherapy Planning and Image Segmentation

One of the most robust application areas for AI lies in image segmentation and radiotherapy planning. Deep learning–based autocontouring systems have consistently demonstrated substantial reductions in planning time while achieving contour accuracy comparable to expert manual delineations. Multi-institutional studies report time savings of up to 50–70% for organs-at-risk and target volumes, particularly in head-and-neck and prostate cancer workflows (35).

Commercial implementations, including systems integrated into clinical radiotherapy platforms and cloud-based solutions (e.g., Microsoft-supported research collaborations), illustrate how AI can shift professional effort from repetitive manual tasks toward quality assurance and clinical decision-making. Nonetheless, contouring errors - especially in anatomically complex or post-operative cases - remain clinically relevant, underscoring the continued need for expert review (36).

5.4 Evidence Quality, Limitations, and Generalizability

Despite promising results, the current evidence base for AI in radiology remains heterogeneous. Many studies are retrospective, single-center, or enriched with high disease prevalence, limiting external validity. Systematic reviews highlight frequent shortcomings in study design, including limited reporting on failure modes, insufficient subgroup analysis, and lack of prospective outcome data. (37)
Moreover, performance degradation after deployment—due to dataset shift, protocol changes, or evolving disease patterns—has been documented, emphasizing that AI systems require continuous monitoring and recalibration rather than one-time approval. (30)

5.5 Human Oversight and Clinical Integration

Across all application domains, a consistent conclusion emerges: AI systems deliver the greatest benefit when deployed as decision-support tools embedded within clinical workflows, not as replacements for human expertise. Human–AI collaboration has been shown to outperform either alone in multiple diagnostic tasks, particularly when AI outputs are presented transparently and radiologists are trained to interpret algorithmic confidence and limitations. (38)
Accordingly, regulatory authorities and professional societies converge on the principle that accountability remains with the physician, and that AI systems must be auditable, explainable to an appropriate degree, and aligned with clinical responsibility frameworks (29, 39).

5.6 Summary

AI applications in radiology already demonstrate tangible gains in efficiency, standardization, and task-specific diagnostic performance. Radiotherapy planning, chest X-ray interpretation, and reporting support represent particularly mature use cases. However, current evidence also highlights limitations related to generalizability, dataset bias, and long-term performance stability. Sustainable clinical value therefore depends not only on algorithmic accuracy but on context-aware implementation, continuous evaluation, and robust human oversight.

6. Generative AI: Support Rather Than Replacement

For several years, the discourse surrounding artificial intelligence in radiology was dominated by predictions of professional displacement. This narrative was epitomized by Geoffrey Hinton's widely cited statement in 2016 suggesting that "we should stop training radiologists," reflecting the belief that image recognition tasks would soon be fully automated by deep learning systems. Nearly a decade later, empirical evidence and clinical experience have demonstrated the opposite: radiologists are not being replaced but are increasingly integrating AI - particularly generative AI - into their workflows as a supportive technology.

6.1 From Automation Anxiety to Augmentation Reality

Early concerns about replacement were largely driven by narrow task-based benchmarks in image classification, where AI systems matched or exceeded human performance under controlled conditions. However, real-world radiology encompasses far more than image recognition, including clinical reasoning, contextual interpretation, communication, quality assurance, and interdisciplinary coordination. Subsequent analyses have emphasized that these broader competencies are not amenable to full automation and instead benefit from human – AI collaboration (30, 34).

Generative AI systems—based on large language models (LLMs) and multimodal architectures—mark a conceptual shift from diagnostic automation toward cognitive and administrative support. Rather than issuing autonomous diagnoses, these systems assist with report drafting, clinical summarization, protocol suggestions, and information retrieval, thereby reducing cognitive load and time spent on non-interpretative tasks (29).

6.2 Generative AI in Reporting and Documentation

One of the most immediate applications of generative AI in radiology is report generation and structuring. LLM-based systems can draft preliminary reports from structured inputs, prior examinations, and clinical context, which are then reviewed, edited, and finalized by radiologists.

Early studies indicate that such tools can reduce reporting time and improve consistency, particularly for standardized examinations, while maintaining physician oversight as a safeguard against errors and hallucinations. (34, 40)

Crucially, professional guidance consistently frames generative AI as an assistive technology. The ESR explicitly states that AI-generated text must not replace clinical judgment and that radiologists remain fully accountable for report content and diagnostic conclusions (29). This positioning reflects broader concerns regarding automation bias and underscores the importance of maintaining human responsibility.

6.3 Evidence from Early Clinical Evaluations

Emerging evaluations of generative AI tools in medical documentation suggest that their value lies in workflow efficiency rather than diagnostic autonomy. In a study assessing the use of ChatGPT-like models for radiology-related tasks, performance was found to be variable and highly dependent on prompt structure, task complexity, and clinical supervision, reinforcing that such systems are not reliable as standalone clinical decision-makers. (41)

These findings align with broader healthcare AI literature demonstrating that productivity gains are most pronounced when AI offloads clerical and repetitive tasks, allowing clinicians to reallocate time toward patient interaction, complex decision-making, and quality assurance. (42)

6.4 Professional Roles and Responsibility

The reframing of AI from replacement to support has important implications for professional identity and training. Rather than diminishing the role of radiologists, generative AI amplifies the need for domain expertise, critical oversight, and system literacy. Radiologists must understand AI limitations, recognize erroneous outputs, and contextualize algorithmic suggestions within the clinical picture—skills that cannot be delegated to machines. (30)

Regulatory authorities reinforce this view. The U.S. Food and Drug Administration explicitly emphasizes that AI systems in medicine function as decision-support tools and that accountability remains with the healthcare professional, particularly for adaptive and generative models whose outputs may vary over time. (39)

6.5 Synthesis

Nearly a decade after early predictions of obsolescence, radiology offers a clear example of augmentation rather than replacement. Generative AI is increasingly used to streamline documentation, reporting, and information management, reducing administrative burden while preserving - and in some cases enhancing - clinical quality. The evidence to date supports a model in which generative AI serves as a supportive layer within radiological workflows, contingent on transparency, validation, and continuous human oversight.

7. Regulation, Governance, and Quality Assurance

The rapid expansion of artificial intelligence (AI) in radiology has shifted the discussion from whether AI can perform specific tasks to how such systems can be deployed safely, responsibly, and sustainably in clinical environments. Regulation, governance, and quality assurance (QA) are therefore not administrative add-ons but prerequisites for clinical adoption. Recent literature increasingly emphasizes that the key challenge is lifecycle control of systems that may behave differently across institutions, populations, and time. (7, 8, 10, 12)

7.1 Regulatory frameworks: from permissive innovation to risk-based control

Globally, regulatory approaches to AI in medical imaging are converging toward risk-based models that prioritize patient safety, transparency, and accountability. In Europe, radiology-relevant AI tools are typically considered high-risk systems because they influence diagnostic and therapeutic decisions. The European Society of Radiology (ESR) has emphasized that the upcoming EU AI Act will likely strengthen obligations for human oversight, documentation, transparency, and post-market responsibilities. (18, 19) In parallel, AI products intended for clinical use remain subject to medical device regulations, meaning that radiology departments cannot treat AI as a "software add-on" but must consider it a regulated medical technology.

From a practical standpoint, this regulatory evolution matters because it influences what counts as acceptable evidence. Traditional performance metrics derived from retrospective datasets are increasingly insufficient as a sole basis for adoption, particularly when a system is expected to operate in heterogeneous real-world environments. The regulatory landscape described in the British Journal of Radiology highlights that compliance requirements will continue to expand, not least because AI systems raise unique challenges such as continuous updating, unclear failure modes, and the need for traceable decision pathways. (8)

A further complication arises with the emergence of large language models (LLMs) and generative AI. These systems do not fit neatly into conventional *"locked algorithm"* paradigms. Regulatory approval processes for LLM-based medical devices require additional considerations beyond classical AI, including issues of non-deterministic outputs, susceptibility to hallucinations, and the difficulty of defining stable performance characteristics. (18, 19) For radiology, this is particularly relevant because LLMs may increasingly be used for reporting support, protocol guidance, or clinical summarization —functions that can still affect patient care even if they are not framed as diagnostic classification tools.

7.2 Governance: defining responsibility and preventing accountability gaps

Regulation defines external requirements, but governance determines how an institution operationalizes them. Governance in radiology AI must address three core questions:

- *Who owns the system clinically?*
- *Who is accountable when the system fails?*
- *How is ongoing performance ensured?*

A consistent message across recent sources is that the radiologist remains responsible for the final diagnostic output, even when AI is integrated into the workflow. (7, 8, 10, 12) This principle is not merely a legal formality; it has practical implications. If AI output is treated as authoritative or is integrated in a way that subtly nudges decision-making, then responsibility without control becomes an unsafe model. Therefore, governance must ensure that radiologists retain meaningful oversight and the ability to challenge or disregard AI suggestions.

The need for structured governance is further reinforced by evidence that AI assistance does not benefit all radiologists equally and can sometimes worsen performance when AI is wrong. (7, 8, 10, 12) Such findings undermine simplistic assumptions that AI is uniformly "helpful" and highlight the importance of implementation strategies that explicitly manage human factors, training, and error exposure. Governance must therefore include human–AI interaction considerations, not only technical validation.

A realistic governance model typically requires a multidisciplinary structure. In many institutions, this includes radiology lea-

dership, medical physics, IT and cybersecurity, data protection officers, legal counsel, and clinical stakeholders from high-impact pathways (e.g., emergency medicine, oncology). Governance should define decision rights for procurement, evaluation, deployment, monitoring, and decommissioning. Without such structures, AI adoption risks becoming fragmented, vendor-driven, or dependent on local enthusiasm rather than evidence and oversight.

7.3 Quality assurance as a life-cycle obligation, not a one-time test

A recurring limitation in the AI radiology literature is the mismatch between how AI systems are validated and how they are used. Many AI tools demonstrate performance in retrospective datasets but are deployed into workflows with different prevalence, imaging protocols, patient populations, and operational constraints. The *"emperor has few clothes"* critique captures this gap sharply: AI systems may appear impressive in controlled settings, yet evidence for efficiency gains and robust real-world impact remains limited. (18, 19) This critique is not anti-technology; it is a reminder that clinical value depends on implementation and sustained performance.

Quality assurance for radiology AI must therefore be conceptualized as continuous. Testing processes described in the medical physics literature emphasize that evaluation should cover not only algorithm performance but also integration, failure handling, and reproducibility. (7, 8, 10) In practice, QA needs at least three layers:

1. Pre-deployment validation (local acceptance testing)

Before routine use, AI systems should be tested on local data that reflect the institution's scanners, protocols, patient mix, and clinical prevalence. This step helps identify dataset shift early. It also provides baseline metrics against which future drift can be detected. A crucial governance decision is whether the system is used as a second reader, triage tool, or quantification aid—each use case implies different risk profiles and QA requirements.

2. Deployment monitoring (performance surveillance)

- Post-market surveillance is frequently mentioned as essential, partly because prospective randomized trials are often too resource-intensive and too slow for rapidly evolving software. (11) Monitoring should include:

- basic performance indicators (e.g., sensitivity proxies, false-positive rates where measurable)
- workflow metrics (time-to-report, case prioritization effects)
- discrepancy and incident tracking
- user feedback (radiologist trust, perceived failure modes)
- importantly, monitoring should not rely solely on vendor dashboards. Institutions need independent capacity to detect unexpected behavior, particularly in high-risk pathways.

3. Periodic re-evaluation and controlled updating

- AI systems may degrade over time due to changes in scanners, reconstruction algorithms, patient demographics, or clinical practice. Additionally, software updates may change performance characteristics. Governance must ensure that updates are treated as clinically relevant events requiring revalidation. This becomes more complex with adaptive AI systems and even more so with LLM-based components, where outputs may vary and *"version stability"* can be difficult to define. (18, 19)

7.4 Managing risk: from technical errors to socio-technical failure modes

Traditional medical device QA often focuses on technical accuracy and hardware reliability. AI introduces new categories of risk, including sociotechnical failures where harm results from the interaction between humans, software, and workflow.

One prominent example is automation bias—overreliance on automated suggestions. Evidence indicates that incorrect AI predictions can adversely affect radiologist performance on aggregate and for specific tasks. (7, 8, 10, 12) This suggests that AI errors are not merely additive but can propagate through human decision-making. A governance and QA system must therefore consider not only *"how often AI is wrong,"* but also *"what happens when AI is wrong."*

Risk management must also address the possibility of miscalibration, particularly when AI is deployed in populations with different disease prevalence. A system trained in one setting may produce misleading probability estimates in another, affecting both radiologist interpretation and clinical decision-making. (11) This supports the argument that local validation and calibration checks are not optional extras but essential safety steps.

7.5 Documentation, transparency, and auditability

From a regulatory and clinical governance perspective, documentation is a practical necessity. Radiology departments must be able to answer basic questions:

- *What does the AI do?*
- *On which data was it trained?*
- *How is it intended to be used?*
- *What are known limitations?*
- *What performance has been demonstrated locally?*
- *What happens when the AI output conflicts with radiologist judgment?*

In Europe, the move toward stronger regulatory oversight will likely increase expectations for documentation, audit trails, and transparency. (7, 8, 10) Legal analyses also emphasize that unclear documentation and undefined responsibility boundaries create liability risks. (12, 13) For quality assurance, documentation supports reproducibility and learning: when an AI-related incident occurs, it must be possible to reconstruct what the system output was, how it was displayed, and how the clinician responded.

7.6 A pragmatic synthesis: what "good governance" looks like in radiology

Taken together, recent evidence and expert guidance suggest that successful AI deployment in radiology depends less on single performance numbers and more on robust governance and QA. Regulation sets minimum standards, but departments must translate them into operational practice. Testing must be local and life-cycle oriented, monitoring must be continuous, and human oversight must be meaningful rather than symbolic. (7, 8, 10, 12)

A conservative and realistic conclusion is that AI in radiology is best treated as a high-impact clinical technology that requires the same discipline as any other medical device —while acknowledging that its risks are often less visible and more workflow-dependent. Radiology departments that invest early in governance structures, QA processes, and post-deployment monitoring are more likely to realize sustainable benefits and less likely to experience harmful surprises. The central goal should not be rapid adoption, but safe and accountable integration into clinical care.

8. Limit of Current Systems: "Common Sense" and Explainability

Despite measurable progress in narrow radiological tasks, current AI systems remain fundamentally limited in ways that are clinically relevant and often underestimated in implementation discussions. These limitations are not primarily about raw pattern recognition, where deep learning has demonstrated strong performance in many settings, but about robustness, contextual reasoning, and the ability to behave safely when confronted with uncertainty, atypical presentations, or shifting clinical environments. In other words, contemporary AI may appear competent within well-defined test conditions, yet still lack the "common sense" required for reliable operation in real-world radiology.

8.1 "Common sense" in radiology: more than image classification

Radiological interpretation is not simply a matter of detecting abnormalities. It is an integrative cognitive process that combines imaging findings with clinical context, prior examinations, pre-test probability, and downstream consequences. Radiologists routinely perform tasks that are difficult to formalize: weighing differential diagnoses, recognizing when an image is technically inadequate, identifying incidental findings that matter (and those that do not), and tailoring recommendations to patient-specific circumstances.

Current AI systems typically do not possess this form of contextual reasoning. They excel at specific tasks under predefined conditions but struggle when the clinical question

changes, when imaging protocols differ, or when unexpected confounders occur. This limitation is particularly important because radiology is full of "edge cases": post-operative anatomy, mixed pathologies, rare diseases, artifacts, and incomplete clinical information. A model that performs well on average may still fail in precisely the cases where radiologists add the most value.

The gap between narrow task performance and real-world clinical utility contributes to a broader sense of "post-hype realism." A critical appraisal has argued that many AI tools currently add complexity without proportionate efficiency gains, particularly when they are layered onto existing workflows rather than replacing clearly defined tasks. (11) In such scenarios, radiologists still must read the entire case, verify AI outputs, and manage exceptions—meaning that the AI does not remove work but can create additional cognitive load.

8.2 Robustness and generalizability: the persistent problem of dataset shift

A central technical and clinical limitation is generalizability. Many AI systems are trained and validated on datasets that do not represent the full heterogeneity of clinical practice. Differences in scanners, reconstruction algorithms, acquisition parameters, patient demographics, disease prevalence, and institutional workflows can produce dataset shift that degrades performance.

In radiology, such shifts are not rare; they are routine. A system validated in a tertiary academic center may behave differently in a community hospital. A model trained on one vendor's imaging data may fail silently on another. Even within the same institution, protocol updates or software upgrades can change image appearance enough to influence model outputs. These effects are difficult to predict from retrospective validation alone, reinforcing the need for ongoing monitoring and post-market surveillance. (7, 10, 11)

A related concern is that performance metrics reported in studies often mask clinically relevant failure modes. High AUC values can coexist with systematic errors in subgroups or with poor calibration in real-world prevalence settings. The clinical risk is not only that the AI is imperfect, but that its errors may not be obvious to users—particularly when the system presents confident outputs without reliable uncertainty information.

8.3 Human factors: why AI errors are not neutral

In radiology, AI errors are not necessarily independent of human performance. Reader studies have shown that AI assistance can have heterogeneous effects across radiologists, and that incorrect AI predictions can adversely influence radiologist performance on aggregated tasks and on specific pathologies. (1) This is a crucial point: AI is not merely an additional opinion, but a cognitive input that can bias interpretation.

Such effects align with well-known human factors phenomena, including automation bias and authority bias. When AI is presented as *"smart"* or *"validated"*, users may overweight its suggestions, especially under time pressure or in ambiguous cases. The practical implication is that the safety profile of AI is not determined solely by its standalone accuracy, but by the interaction between AI outputs, human decision-making, and workflow design. This makes explainability, calibration, and appropriate user training more than academic concerns; they become patient safety issues.

A conservative interpretation is therefore warranted: even high-performing AI systems can reduce overall diagnostic quality if they are integrated in a way that increases over-reliance or disrupts radiologists' normal verification strategies. The goal of implementation should not be to maximize AI visibility, but to ensure that AI outputs are presented in ways that support sound clinical judgment.

8.4 Explainability: promises, limits, and practical relevance

Explainability is frequently proposed as a solution to trust and safety concerns. In principle, explainable AI should allow users to understand why a system produced a given output, identify when it is likely to be wrong, and maintain meaningful oversight. In practice, however, explainability remains limited and sometimes misunderstood.

Many commonly used explainability methods in imaging (e.g., saliency maps or heatmaps) can provide visually appealing overlays but

do not necessarily correspond to clinically meaningful reasoning. They may highlight regions correlated with model predictions without proving causal understanding. Furthermore, explanations can create a false sense of security: users may interpret an explanation as evidence of correctness, even when the model is wrong.

From a clinical perspective, the most useful "explainability" may not be a visual overlay, but robust transparency about model scope, limitations, and uncertainty.
This includes:
- *what the model was trained on,*
- *which populations and scanners were represented,*
- *what kinds of errors are common,*
- *and how performance changes with prevalence.*

These elements align closely with governance and quality assurance requirements, including documentation, auditability, and monitoring. (7, 8, 10, 12)

8.5 LLMs and generative systems: a new category of limitations

Large language models introduce additional constraints beyond conventional diagnostic AI. While LLMs can support radiology through report drafting, summarization, protocol guidance, or structured reporting, they are prone to hallucinations, non-deterministic outputs, and sensitivity to prompt wording.
These characteristics make stable validation difficult and raise questions about how such systems can be regulated as me-dical devices. (9)

The regulatory and methodological challenges of LLMs are not theoretical. If an LLM generates a plausible but incorrect statement in a report draft, the error may be difficult to detect—particularly in high-volume settings. Moreover, the output may appear confident and fluent, increasing the risk of overtrust.

A realistic approach is therefore to treat generative AI as a supportive layer that requires strict governance, constrained use cases, and careful QA, rather than as an autonomous clinical agent. (7, 8, 10, 12)

8.6 Summary: why limitations matter for safe clinical adoption

The limitations of current AI systems in radiology are not best described as "AI is not good enough," but rather as "AI is good at some things, yet unreliable in ways that matter clinically." The gap between narrow performance and real-world robustness, the absence of common-sense contextual reasoning, the challenges of dataset shift, and the complexities of human–AI interaction all argue for a cautious approach.

Explainability may contribute to safer deployment, but it should not be treated as a universal remedy. Instead, safe adoption requires a combination of conservative use-case selection, strong governance, local validation, continuous monitoring, and training that addresses human factors. (7, 8, 10, 12) In this framing, radiologists remain central—not because AI is ineffective, but because current systems lack the broader clinical reasoning and responsibility that define radiological practice.

9. Outlook: Agentive Systems and Division of Labor

The near-term future of AI in radiology is unlikely to be defined by autonomous *"replacement"* of radiologists, but rather by a gradual restructuring of workflows and responsibilities. As AI tools become more capable and increasingly integrated into clinical systems, the central question shifts from *"Can AI interpret images?"* to *"Which parts of radiological work should be delegated to machines, and under what governance conditions?"* This outlook requires a pragmatic concept of division of labor: assigning tasks to AI where it is demonstrably reliable, measurable, and safe, while preserving human responsibility for synthesis, context, and final decision-making. (7, 8, 10, 12)

9.1 From isolated tools to orchestrated work-flows

Most current radiology AI systems are narrow applications—detection algorithms, quantification tools, or triage aids. These tools often operate as add-ons to existing workflows. A consistent critique is that such add-on deployment may increase complexity

without delivering proportional efficiency gains, because radiologists must still perform full reads and maintain oversight. (11) The next stage of development is therefore expected to focus less on adding more "point solutions" and more on integrating AI into coordinated workflows that remove clearly defined burdens from radiologists.

This evolution is sometimes described as a shift toward "agentive systems": software components that can execute multi-step processes across systems rather than providing a single prediction. In radiology, an agentive workflow might automatically retrieve priors, align follow-up studies, compare measurements longitudinally, detect discrepancies, draft structured summaries, and surface cases that require urgent attention. While this vision is technologically plausible, it raises immediate governance questions. If an agent coordinates multiple tasks, errors may propagate through the workflow, and accountability can become diffuse unless responsibility boundaries are explicitly defined. (7, 8, 10, 12)

A conservative outlook therefore recognizes that the future is not merely *"more AI,"* but *"more interconnected AI,"* which increases both potential benefits and potential failure modes.

9.2 A realistic division of labor: what AI can do well

A practical division of labor should prioritize tasks that are (a) repetitive, (b) time-consuming, (c) measurable, and (d) less dependent on nuanced clinical context. In radiology, this often includes:

a) Image quality and protocol support: identifying incomplete acquisitions, suggesting repeat sequences, and flagging technical limitations.
b) Quantification and measurement: volumetry, lesion segmentation, and longitudinal change tracking—particularly where manual measurement is inconsistent or burdensome.
c) Prior comparison and follow-up tracking: automatically aligning prior studies, highlighting interval changes, and ensuring relevant comparisons are not missed.
d) Workflow triage: flagging potentially urgent findings to reduce time-to-action in high-risk pathways.
e) Structured reporting assistance: populating templates, ensuring completeness, and reducing clerical burden.
f) Administrative automation: coding support, worklist management, and report distribution tasks.

These areas align with the broader argument that radiology's sustainability depends not only on marginal gains in diagnostic accuracy, but on meaningful reductions in workload and improved system efficiency. (11) Importantly, such applications also tend to be easier to validate and monitor than complex "end-to-end diagnostic reasoning" systems.

9.3 What remains distinctly human: synthesis, accountability, and clinical judgement

Even as AI takes on a larger share of measurable tasks, several responsibilities remain inherently human, at least for the foreseeable future:

a) Clinical synthesis and prioritization
Radiologists interpret imaging in the context of incomplete information, competing differentials, and variable clinical relevance. This includes deciding what matters, what can be ignored, and what requires immediate escalation.

b) Handling ambiguity and rare events
Radiology is characterized by exceptions, artifacts, and unusual combinations of findings. AI may perform well on common patterns but is less reliable in atypical situations, particularly under dataset shift. (11)

c) Ethical and professional accountability
Clinical responsibility cannot be delegated to a model. Legal analyses emphasize that accountability remains with the medical professional and the institution, even when AI is involved. (12)

d) Communication and consultation
Discussing findings with clinicians and patients, resolving contradictions, and translating imaging into actionable recommendations remain core radiologist functions. This consultative role is difficult to automate safely.
This division of labor reflects a broader principle: radiology is not only image interpretation but a clinical service embedded in

patient pathways. AI can support this service, but cannot replace its professional responsibility structure.

9.4 Agentive systems and governance: avoiding accountability gaps

As AI becomes more agent-like—executing sequences of actions rather than producing isolated outputs—governance must evolve accordingly. Traditional approval and QA models assume relatively static systems with predictable behavior. Agentive workflows may involve dynamic interactions between multiple software components, increasing complexity and reducing transparency.

This creates a risk of "accountability gaps", where no individual can fully explain why a certain workflow produced a given output. The regulatory and governance literature increasingly emphasizes life-cycle control, documentation, human oversight, and monitoring as safeguards against such gaps. (7, 8, 10, 12) The challenge is to ensure that oversight remains meaningful. *"Human-in-the-loop"* must not become a symbolic phrase that simply transfers responsibility to radiologists without giving them control or visibility into system behavior.

In practice, safe governance of agentive workflows likely requires:
a) clear definition of AI scope and intended use,
b) audit trails for key decisions and outputs,
c) controlled updates and revalidation,
d) incident reporting and corrective action processes,
e) and explicit fallback strategies when AI outputs are unavailable or inconsistent.

These elements mirror medical device safety principles but must be adapted to software that may change more rapidly and interact more broadly with clinical systems. (7, 10)

9.5 A conservative outlook: incremental transformation rather than disruption

A realistic future for radiology is one of incremental transformation rather than sudden disruption. AI will likely be adopted where it solves concrete problems—reducing repetitive workload, improving consistency of quantification, and supporting workflow prioritization—while radiologists remain responsible for interpretation, integration, and patient-centered decision-making.

The most successful implementations will likely be those that treat AI not as a replacement technology, but as a workforce multiplier under strict governance. This approach aligns with the post-hype phase of radiology AI, where the emphasis shifts from performance claims to evidence, safety, and sustainable value creation. (7, 8, 10, 12) In this framing, agentive systems may become valuable tools, but only if their integration is guided by conservative governance and rigorous QA rather than by technological enthusiasm alone. (7, 10, 12)

10. Practical Checklist: Governance & Safe Implementation

Successful implementation of AI in radiology depends less on "algorithmic performance in principle" and more on disciplined governance, local validation, and continuous quality assurance. Recent literature consistently highlights that real-world adoption is constrained by regulatory obligations, human factors, workflow complexity, and limited economic evidence. (7, 8, 10, 12) In this setting, a pragmatic checklist can support departments in moving from interest-driven adoption to safe, auditable, and clinically meaningful deployment.

The following checklist is designed for radiology departments planning to introduce AI systems for clinical use. It focuses on high-risk decision support tools but is applicable to most AI applications, including workflow and reporting support. It reflects core principles of staged clinical evaluation (5), health-economic transparency (6), regulatory alignment (7, 8, 12), and practical testing processes in radiology environments. (10)

10.1 Governance and accountability

a) Define the clinical owner (*"responsible physician"*)

b) Named radiologist accountable for clinical oversight and intended use.

c) Define decision rights and escalation pathways

d) Who can approve deployment, pause use, or decommission the system?

e) Clarify responsibility boundaries

f) AI output is advisory; final responsibility remains with clinicians. (12)

g) Establish a multidisciplinary oversight group

h) Radiology leadership, IT/security, medical physics, legal/compliance, data protection, key referrers.

10.2 Use-case definition and risk classification

a) Specify the intended use precisely

b) Detection, triage, quantification, second reader, reporting support, etc.

c) Define patient population and clinical pathway

d) Emergency, screening, oncology follow-up, MS monitoring, etc.

e) Identify failure modes with clinical risk assessment

f) False negatives vs false positives; consequences and mitigation.

g) Confirm regulatory status and labeling

h) CE marking / regulatory clearance for the intended use. (7, 8)

10.3 Data governance, privacy, and cybersecurity

a) Confirm legal basis for data processing

b) Local privacy laws, institutional approvals, contracts.

c) Ensure secure integration

d) Network segmentation, authentication, logging, vulnerability management.

e) Clarify data flows

f) What leaves the hospital? Cloud processing? Storage duration? (12)

g) Vendor transparency requirements

h) Training data description, versioning, update policies, audit support. (7, 8)

10.4 Local validation before deployment ("acceptance testing")

a) Test on local representative cases

b) Scanner types, protocols, prevalence, demographics.

c) Define performance metrics and thresholds upfront

d) Sensitivity/specificity proxies, false-positive burden, time impact.

e) Compare against current standard of care

f) Ensure AI adds measurable value rather than complexity overhead. (11)

g) Validate workflow behavior

h) Where AI output appears, how it is displayed, how it is acted upon. (10)

10.5 Workflow integration and human factors

a) Define when and how radiologists see AI results

b) Early triage vs after initial read; avoid over-anchoring.

c) Train users (radiologists, technologists, clinicians)

d) Intended use, limitations, known failure modes, escalation rules. (5)

e) Address automation bias explicitly

f) Encourage independent verification, especially in ambiguous cases. (1, 11)

g) Provide a clear "AI off" fallback mode

h) Ensure continuity of care if AI fails or is paused.

10.6 Post-deployment monitoring and quality assurance

a) Establish continuous performance monitoring

b) Drift detection, subgroup issues, unexpected false positives/negatives. (7, 10)

c) Implement incident reporting and review

d) Near misses, discrepancies, adverse events linked to AI output.

e) Monitor workflow impact

f) Turnaround time, prioritization effects, radiologist workload.

g) Revalidate after major changes

h) Scanner upgrades, protocol changes, software updates, vendor model updates. (7, 10)

10.7 Economic and operational evaluation

a) Document implementation costs

b) Licenses, infrastructure, staff time, training, integration.

c) Define expected value outcomes

d) Efficiency gains, reduced rereads, improved quantification consistency.

e) Perform transparent health-economic assessment where feasible

f) Use CHEERS-AI principles for reporting and interpretation. (6)

10.8 Documentation and auditability

a) Maintain an AI system dossier

b) Intended use, validation results, version history, known limitations.

c) Ensure audit trails for AI outputs

d) Output storage, timestamps, user interaction logs where possible. (7, 8)

e) Define review cycles

f) Quarterly/annual governance review; decision to continue, adjust, or stop.

10.9 Checklist summary

AI implementation in radiology should be treated as a controlled clinical intervention rather than a plug-in technology. A conservative governance model emphasizes clear accountability, precise use-case definition, local validation, continuous monitoring, and structured response to drift or failure. (18, 19) This approach supports both patient safety and sustainable clinical value, while reducing the risk that AI adoption becomes driven by expectations rather than evidence. (5, 6, 12)

Discussion

Artificial intelligence has moved from a largely experimental technology to a set of clinically deployed tools that increasingly influence radiological workflows. The present article intentionally adopts a conservative and practice-oriented perspective: rather than focusing on technological potential alone, it integrates current evidence, implementation frameworks, radiation safety culture, and regulatory developments to assess where AI already contributes meaningful value and where limitations remain. Across the reviewed sources, a clear trend emerges: the field has transitioned from early enthusiasm and disruption narratives to methodological realism and a stronger emphasis on governance, quality assurance, and sustained clinical responsibility. (11)

Clinical value and the role of radiology in modern healthcare

Radiology's contribution to clinical care is best understood as multidimensional. It includes diagnostic accuracy, timely and actionable interpretation, consultation, multidisciplinary integration, patient communication, and stewardship of appropriate imaging. (14, 15, 17, 20, 22) Importantly, radiology's value is not always captured by traditional productivity metrics such as report volume or turnaround time. Instead, its impact is often indirect, distributed across clinical decisions and patient pathways. This makes "value" harder to measure, but not less real.

The value-based healthcare perspective reinforces that radiology cannot be reduced to a commodity service. Radiologists create value when imaging results are integrated into clinical reasoning and translated into management-relevant conclusions. (14, 15, 22) In this context, the radiologist's consultative role becomes central, especially in complex cases where interpretation depends on context and where imaging findings must be weighed against differential diagnoses, risks, and downstream consequences. (17, 20)

Patient-centered value also deserves explicit attention. Surveys and feedback-focused initiatives suggest that patients and referring

clinicians increasingly evaluate radiology not only by technical quality, but by clarity, communication, responsiveness, and trust. (18, 19) This has implications for reporting style, accessibility of results, and radiology's visibility within the healthcare system. (16, 23)

Evidence quality: progress, heterogeneity, and persistent gaps

While AI systems have demonstrated strong performance in specific tasks, the evidence base remains heterogeneous. Many studies are retrospective, use enriched datasets, or evaluate performance under controlled reader-study conditions that do not fully reflect routine clinical complexity. (1–4) Prospective outcome-driven trials remain relatively rare, and external validation across diverse institutions and patient populations is still limited. (2, 3, 11)

A critical insight from human–AI interaction research is that AI assistance does not uniformly improve performance. The effect of AI depends on the task, the reader, the clinical setting, and the error profile of the system. (1) Incorrect AI outputs can negatively influence radiologists, illustrating that AI errors are not neutral but can propagate through cognitive bias and workflow pressure. (1, 11) These findings challenge simplistic narratives of universal benefit and support a cautious approach to deployment, particularly in high-risk pathways.

Methodological frameworks such as DECIDE-AI and guidance on clinical evaluation highlight the need for staged evidence generation, transparency, and context-aware reporting. (4, 5) Health-economic standards such as CHEERS-AI further emphasize that claims of value must include implementation costs, workflow effects, and real-world constraints, rather than relying on speculative efficiency arguments. (6) Taken together, these frameworks reflect a broader shift: AI in radiology must be evaluated as a clinical intervention embedded in complex systems, not as a standalone algorithm.

Radiation protection and AI: complementary safety cultures

The inclusion of radiation protection as a cultural concept (technology, behavior, organization) provides an important parallel to AI governance. Radiation safety has long been recognized as a socio-technical challenge: technology enables dose optimization, but outcomes depend on training, behavior, organizational structures, and continuous monitoring. (24 – 28) This logic applies directly to AI. As with radiation protection, safe AI adoption requires not only technical tools but also professional responsibility, institutional processes, and a learning culture.

This perspective supports the argument that AI governance should be integrated into existing safety and quality infrastructures rather than treated as a separate "digital innovation" track. Radiology departments already have experience managing complex technologies with invisible risks; AI extends this responsibility into the domain of software-driven decision support.

Generative AI: meaningful support, but not autonomy

Generative AI and large language models introduce new opportunities and risks. Their value may lie primarily in administrative and cognitive support, such as report drafting, structured documentation, summarization, and information retrieval. (29, 34, 40) However, these systems are prone to hallucinations, instability across prompts, and outputs that may appear plausible while being incorrect. (9, 41) Therefore, their safe use requires constrained use cases, strict oversight, and careful workflow design.

The emerging consensus across regulatory and professional discussions is that generative AI should be positioned as support rather than replacement. (29, 39) This framing aligns with the clinical reality that radiology depends on contextual reasoning, accountability, and communication—elements that current AI systems cannot reliably replicate.

Regulation, governance, and quality assurance as prerequisites

Regulatory developments are increasingly shaping AI adoption in radiology. In Europe, radiology AI tools are typically treated as high-risk systems, implying stronger requirements for documentation, transparency, human oversight, and life-cycle risk management. (7, 8, 12) These requirements are not merely legal constraints; they define the minimum conditions for responsible clinical use.

A key challenge is that AI performance is not static. Dataset shift, protocol changes, population differences, and software updates can degrade performance over time. (10, 11) This makes quality assurance a continuous obligation rather than a one-time validation step.

Testing processes proposed for clinical environments emphasize local acceptance testing, monitoring, and controlled updating. (10) Without these measures, even well-performing systems may become unreliable in practice.

The discussion of governance also highlights a central ethical and professional principle: radiologists remain accountable for clinical decisions, even when AI is involved. (12) This requires that AI tools are implemented in ways that preserve meaningful human oversight, rather than shifting responsibility without control. In practice, robust governance structures, auditability, and incident management processes are necessary to prevent *"accountability gaps,"* particularly as AI becomes more integrated into multi-step workflows. (7, 10)

Limits of current systems and realistic expectations

The limitations of current AI systems are best understood as limitations in robustness, common-sense reasoning, and clinical generalization rather than limitations in narrow pattern recognition. (11) AI can be strong within defined boundaries but remains vulnerable to atypical cases, confounders, and shifts in clinical reality. Explainability methods may improve transparency, but they do not fully solve the deeper problem of contextual reasoning and safe behavior under uncertainty. (9, 11)

Therefore, a realistic near-term outlook is incremental transformation rather than disruption. AI will likely deliver value where tasks are repetitive, measurable, and well-defined—such as quantification, segmentation, triage support, and structured reporting—while radiologists remain essential for synthesis, contextual interpretation, consultation, and responsibility. (1, 7, 10, 11)

Implications for practice

The practical checklist provided in this article translates these insights into implementable steps. It emphasizes governance, use-case definition, local validation, monitoring, documentation, and health-economic evaluation. (5 – 8, 10, 12) Importantly, such checklists should not be seen as bureaucratic burdens but as safety instruments comparable to established radiology QA practices.

A conservative implementation strategy does not slow innovation unnecessarily; rather, it protects patients and radiologists from preventable failures and supports sustainable adoption. The ultimate goal is not rapid deployment, but clinically meaningful integration with demonstrable benefit.

Conclusion

Artificial intelligence is increasingly becoming a practical component of radiology, but its clinical value depends less on headline performance metrics and more on evidence-based implementation, robust governance, and sustained human oversight. Current AI systems can improve efficiency and support task-specific performance, yet limitations in generalizability, human–AI interaction effects, and real-world robustness remain substantial. Radiology therefore enters a post-hype phase in which responsible adoption requires methodological rigor, continuous quality assurance, and alignment with evolving regulatory frameworks. Rather than replacing radiologists, AI is best understood as a supportive technology that can strengthen radiology's clinical role—provided that accountability, safety culture, and patient-centered value remain the guiding principles.

Correspondence to:
Frank Mosler
Department of Diagnostic, Interventional and Pediatric Radiology, Inselspital Bern, University of Bern, Bern, Switzerland
INSELSPITAL
UNIVERSITÄTSSPITAL BERN | Universitätsinstitut für Diagnostische, Interventionelle und Pädiatrische Radiologie
HÔPITAL UNIVERSITAIRE DE BERNE

Declarations

Consent for publication: The author clarifies that written informed consent was obtained and the anonymity of the patient was ensured. This study submitted to Swiss J. Rad. Nucl. Med. has been conducted in accordance with the Declaration of Helsinki and according to requirements of all applicable local and international standards. All authors contributed to the conception and design of the manuscript, participated in drafting and revising the content critically for important intel-lectual input, and approved the final version for publication. Each author agrees to be accountable for all aspects of the work, ensuring its accuracy and integrity.

Competing interests: None.

Funding: No funding was required for this study.

Conflict of interest:

The authors declare that there were no conflicts of interest within the meaning of the recommendations of the International Committee of Medical Journal Editors when the article was written.

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 Swiss J. Radiol. Nucl. Med. and/or the editor(s). Swiss J. Radiol. Nucl. Med. 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.

License Policy:

SJORANM-LinkedIn:

Check out our journal's LinkedIn profile with over 11K registered followers from the Radiologic & Nuclear Medicine Imaging field.

References

1. Yu F, Moehring A, Banerjee O, et al. Heterogeneity and predictors of the effects of artificial intelligence assistance on radiologists. Nat Med. 2024. https://doi.org/10.1038/s41591-024-02850-w

2. van Winkel SL, Peters J, Janssen N, et al. Artificial intelligence as an independent second reader for breast cancer screening: a retrospective cohort study. Lancet Digit Health. 2025. https://doi.org/10.1016/j.landig.2025.100882

3. Barnett M, Wang D, Beadnall HN, et al. Real-world clinical validation of an artificial intelligence system for MRI monitoring in multiple sclerosis. npj Digit Med. 2023;6:197. https://doi.org/10.1038/s41746-023-00940-6

4. Park SH, Han K, Jang HY, et al. Methods for clinical evaluation of artificial intelligence algorithms for medical diagnosis. Radiology. 2023;307(2):e220182. https://doi.org/10.1148/radiol.220182

5. Vasey B, Nagendran M, Campbell B, et al; DECIDE-AI Expert Group. Reporting guideline for the early-stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI. Nat Med. 2022;28(5):924–933. https://doi.org/10.1038/s41591-022-01772-9

6. Elvidge J, Hawksworth C, Saygın Avşar T, et al. Consolidated Health Economic Evaluation Reporting Standards for interventions that use artificial intelligence (CHEERS-AI). Value Health. 2024. https://doi.org/10.1016/j.jval.2024.05.006

7. Kotter E, Brady AP, McEnery T, et al; ESR AI Working Group. Guiding artificial intelligence in radiology: ESR recommendations for effective implementation of the European AI Act. Insights Imaging. 2025;16:19. https://doi.org/10.1186/s13244-025-01905-x

8. Hill DLG. Artificial intelligence in imaging: the regulatory landscape. Br J Radiol. 2024;97(11

57):20240002. https://doi.org/10.1093/bjr/tqae002

9. Park SH. Challenges and additional considerations for regulatory approval of large language models as medical devices. Radiology. 2024;312(1):e241703. https://doi.org/10.1148/radiol.241703

10. Ketola JHJ, Inkinen SI, Mäkelä T, et al. Testing processes for artificial intelligence applications in radiology practice. Phys Med. 2024;118:104842. https://doi.org/10.1016/j.ejmp.2024.104842

11. Huisman M, van Ginneken B, Harvey H. The emperor has few clothes: a realistic appraisal of current artificial intelligence in radiology. Eur Radiol. 2024. https://doi.org/10.1007/s00330-024-10664-0

12. Morales Santos Á, Lojo Lendoiro S, Rovira Cañellas M, Valdés Solís P. Legal regulation of artificial intelligence in the European Union: a practical guide for radiologists. Radiologia (Engl Ed). 2024. https://doi.org/10.1016/j.rxeng.2023.11.011

13. Garni et al. The Regulatory Needs for Radiation Protection Devices based upon Artificial Intelligence: State task or leave unregulated?. (2024). Swiss Journal of Radiology and Nuclear Medicine, 5(1), 5. https://doi.org/10.59667/sjoranm.v5i1.11

14. European Society of Radiology (ESR). Value-based radiology: what is the ESR doing, and what should we do in the future? Insights Imaging. 2021;12:108. https://doi.org/10.1186/s13244-021-01056-9

15. Brady AP, Brink JA, Slavotinek J. Radiology in the Era of Value-Based Healthcare: A Multisociety Perspec-tive. Radiology. 2021. https://doi.org/10.1148/radiol.2020209027

16. European Society of Radiology (ESR). The role of radiologist in the changing world of healthcare: a White Paper of the European Society of Radiology (ESR). Insights Imaging. 2022;13:100. https://doi.org/10.1186/s13244-022-01241-4

17. Ho E. Role of radiology in a multidisciplinary approach to patient care: summary of the ESR International Forum 2022. Insights Imaging. 2023. https://doi.org/10.1186/s13244-023-01377-x

18. Rockall A, Visser JJ, Garcia-Villar C, et al; European Society of Radiology (ESR) Value-based Radiology Subcommittee. Feedback in radiology: Essential tool for improving user experience and providing value-based care. Insights Imaging. 2025. https://doi.org/10.1186/s13244-025-02002-9

19. Fuchsjäger M, Brady AP, et al (ESR Value-Based Radiology Subcommittee). Patient survey of value in relation to radiology: results from a survey of the European Society of Radiology (ESR) value-based radiology subcommittee. Insights Imaging. 2021;12:6. https://doi.org/10.1186/s13244-020-00943-x

20. Iftikhar S, Rahmani S, Zaree OA, et al. The Value of Radiology Consultation: Effort Allocation, Clinical Impact, and Untapped Opportunities. J Am Coll Radiol. 2025. https://doi.org/10.1016/j.jacr.2025.04.006

21. Gutzeit A, Sartoretti E, Reisinger C, et al. Direct communication between radiologists and patients improves the quality of imaging reports. Eur Radiol. 2021. https://doi.org/10.1007/s00330-021-07933-7

22. Brady AP, Brink JA, Slavotinek J. Radiology and value-based healthcare. JAMA. 2020;324(13):1286-1287. https://doi.org/10.1001/jama.2020.14930

23. European Society of Radiology (ESR). The identity and role of the radiologist in 2020: a survey among ESR full radiologist members. Insights Imaging. 2020;11:130. https://doi.org/10.1186/s13244-020-00945-9

24. World Health Organization; International Atomic Energy Agency; International Organization for Medical Physics; International Radiation Protection Association. Enhancing radiation safety culture in health care: guidance for health care providers. World Health Organization; 2024. Published July 10, 2024. Accessed January 17,

2026. https://www.who.int/publications/i/item/9789240091115

25. ICRP, 2007. The 2007 Recommendations of the International Commission on Radiological Protection. ICRP Publication 103. Ann. ICRP 37 (2-4). https://www.icrp.org/publication.asp?id=ICRP%20Publication%20103

26. Smith-Bindman R, Wang Y, Chu P, et al. International variation in radiation dose for computed tomography examinations: prospective cohort study. *BMJ*. 2019;364:k4931. Published 2019 Jan 2. https://doi.org/10.1136/bmj.k4931

27. Smith-Bindman R, Chu P, Wang Y, et al. Comparison of the Effectiveness of Single-Component and Multicomponent Interventions for Reducing Radiation Doses in Patients Undergoing Computed Tomography: A Randomized Clinical Trial. *JAMA Intern Med*. 2020;180(5):666-675. https://doi.org/10.1001/jamainternmed.2020.0064

28. International Atomic Energy Agency, issuing body, et al. Radiation Protection and Safety in Medical Uses of Ionizing Radiation. International Atomic Energy Agency, 2018, https://www.iaea.org/publications/11102/radiation-protection-and-safety-in-medical-uses-of-ionizing-radiation.

29. European Society of Radiology (ESR). What the radiologist should know about artificial intelligence - an ESR white paper. Insights Imaging. 2019;10(1):44. Published 2019 Apr 4. https://doi.org/10.1186/s13244-019-0738-2

30. Kelly CJ, Karthikesalingam A, Suleyman M, Corrado G, King D. Key challenges for delivering clinical impact with artificial intelligence. BMC Med. 2019 Oct 29;17(1):195. https://doi.org/10.1186/s12916-019-1426-2

31. McKinney SM, Sieniek M, Godbole V, et al. International evaluation of an AI system for breast cancer screening. Nature. 2020;577(7788):89-94. https://doi.org/10.1038/s41586-019-1799-6

32. Kiani A, Uyumazturk B, Rajpurkar P, et al. Impact of a deep learning assistant on the histopathologic classification of liver cancer. NPJ Digit Med. 2020;3:23. Published 2020 Feb 26. https://doi.org/10.1038/s41746-020-0232-8

33. Seah JCY, Tang CHM, Buchlak QD, et al. Effect of a comprehensive deep-learning model on the accuracy of chest x-ray interpretation by radiologists: a retrospective, multireader multicase study. Lancet Digit Health. 2021;3(8):e496-e506. https://doi.org/10.1016/S2589-7500(21)00106-0

34. Langlotz CP, Allen B, Erickson BJ, et al. A Roadmap for Foundational Research on Artificial Intelligence in Medical Imaging: From the 2018 NIH/RSNA/ACR/The Academy Workshop. Radiology. 2019;291(3):781-791. https://doi.org/10.1148/radiol.2019190613

35. van Dijk LV, Van den Bosch L, Aljabar P, et al. Improving automatic delineation for head and neck organs at risk by Deep Learning Contouring. Radiother Oncol. 2020;142:115-123. https://doi.org/10.1016/j.radonc.2019.09.022

36. Sharp G, Fritscher KD, Pekar V, et al. Vision 20/20: perspectives on automated image segmentation for radiotherapy. Med Phys. 2014;41(5):050902. https://doi.org/10.1118/1.4871620

37. Liu X, Faes L, Kale AU, et al. A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis. Lancet Digit Health. 2019;1(6):e271-e297. https://doi.org/10.1016/S2589-7500(19)30123-2

38. Ting DSW, Pasquale LR, Peng L, et al. Artificial intelligence and deep learning in ophthalmology. Br J Ophthalmol. 2019;103(2):167-175. https://doi.org/10.1136/bjophthalmol-2018-313173

39. Singh V, Cheng S, Kwan AC, Ebinger J. United States Food and Drug Administration regulation of clinical software in the era of artificial intelligence and machine learning. Mayo Clin Proc Digit Health. 2025;3(3):100231. https://doi.org/10.1016/j.mcpdig.2025.100231

40. Jeblick K, Schachtner B, Dexl J, et al. ChatGPT makes medicine easy to swallow: an exploratory case study on simplified radiology reports. Eur Radiol. 2024;34(5):2817-2825. https://doi.org/10.1007/s00330-023-10213-1

41. Rao A, Kim J, Kamineni M, Pang M, Lie W, Succi MD. Evaluating ChatGPT as an Adjunct for Radiologic Decision-Making. Preprint. medRxiv. 2023;2023.02.02.23285399. Published 2023 Feb 7. https://doi.org/10.1101/2023.02.02.23285399

42. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat Med. 2019;25(1):44-56. https://doi.org/10.1038/s41591-018-0300-7