

Ethical and Governance Challenges of AI in Medical Imaging and Diagnostics: A Systematic Survey and Policy Framework Recommendations

Dulani Athukorala * , Khandakar Ahmed  and Raza Nowrozy 

College of Engineering and Science, Victoria University, Melbourne 3000, Australia;
khandakar.ahmed@vu.edu.au (K.A.); raza.nowrozy@vu.edu.au (R.N.)

* Correspondence: dulani.liyanaathukoralalage@live.vu.edu.au

Abstract

Background/Objectives: Artificial intelligence (AI) is increasingly embedded within diagnostic imaging workflows, reshaping clinical decision-making, health system governance, and regulatory oversight. While technical advances in radiological AI have accelerated, governance mechanisms have struggled to keep pace with issues of bias, transparency, accountability, and lifecycle oversight. This study examines ethical, regulatory, and implementation challenges in AI-enabled diagnostic imaging, building on prior reviews that have often emphasised technical performance by integrating ethical risk domains with governance responses across the AI lifecycle. **Methods:** This study presents a PRISMA-ScR-informed systematic survey of 156 sources, including peer-reviewed publications, regulatory documents, policy reports, and professional guidance materials (2018–2025), synthesised through thematic analysis and lifecycle mapping spanning data acquisition, model development, deployment, monitoring, and continuous learning. **Results:** Drawing on both thematic insights derived from the reviewed literature and established ethical and regulatory frameworks, we propose a literature-derived conceptual ethical-governance framework organised around five pillars: equity and bias mitigation, explainability and transparency, accountability and oversight, privacy-preserving infrastructure, and adaptive regulatory alignment. Although illustrated through the Australian healthcare context, the framework is designed to be transferable to federated and multi-jurisdictional health systems. This review further identifies trust quantification as an underdeveloped but essential dimension of clinical AI governance, emphasising the need to integrate measurable indicators such as calibration, clinician–AI concordance, and patient acceptance into lifecycle-based evaluation. **Conclusions:** By bridging technical, ethical, and policy perspectives, this review proposes a structured conceptual governance framework to support safe, equitable, and trustworthy AI integration in digital health systems.



Academic Editors: Paolo Niccolò Franco and Cesare Maino

Received: 5 May 2026

Revised: 23 June 2026

Accepted: 23 June 2026

Published: 2 July 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

Keywords: artificial intelligence; medical imaging; diagnostic imaging; ethical AI; AI governance; algorithmic bias; explainability; healthcare regulation; digital health; clinical decision support

1. Introduction

In this review, “ethical AI” in diagnostic imaging refers to AI systems that ensure clinical safety and robustness across diverse populations, meaningful transparency for clinician interpretation, documented accountability throughout the AI lifecycle, privacy-preserving data governance, and ongoing monitoring to detect bias, model drift, and

performance degradation. This definition operationalises ethical principles into practical governance requirements for the development, deployment, and oversight of clinical AI systems. In recent years, artificial intelligence (AI), which encompasses machine learning (ML), deep learning (DL), and foundation models, has triggered transformative changes in diagnostic healthcare practices [1–10]. These cutting-edge technologies excel in uncovering diagnostically relevant patterns from high-dimensional medical images, offering levels of accuracy and processing speed that go far beyond traditional approaches. State-of-the-art AI architecture, including convolutional neural networks (CNNs), vision transformers (ViTs), generative adversarial networks (GANs) and large language models (LLMs), have been successfully integrated into clinical imaging workflows, contributing to diverse applications such as anomaly detection, segmentation of anatomical structures, disease classification, radiomic-based feature extraction, individualised risk stratification, and the automation of diagnostic report generation [11–25].

Beyond isolated algorithmic performance improvements, AI-enabled diagnostic imaging systems now function as components of broader digital health infrastructures. These systems interact with electronic health records, clinical decision-support platforms, telehealth networks, and national data repositories. Consequently, governance challenges extend beyond model accuracy to encompass interoperability, data governance, institutional accountability, and public trust within digitally connected health ecosystems. Framing AI in diagnostic imaging as a digital health infrastructure component clarifies why lifecycle governance and system-level oversight are essential.

Crucially, trust in clinical AI systems should not be assumed solely on the basis of technical performance or regulatory approval. Instead, trust must be treated as a measurable and context-dependent outcome, requiring evaluation through empirical indicators such as calibration reliability, clinician–AI concordance, and patient acceptance in real-world settings.

1.1. Scope and Focus of the Systematic Survey

Although the term *medical imaging* encompasses a broad range of diagnostic modalities, including radiology, histopathology, cytology, and molecular imaging, the primary focus of this systematic survey is on AI-enabled radiological imaging workflows. This emphasis reflects the current maturity of AI deployment, regulatory oversight, and clinical integration within radiology, where AI systems are most widely evaluated, validated, and deployed in practice.

While emerging applications in digital pathology and cytology raise important ethical and governance questions, these modalities are considered outside the core analytical scope of this review. Where relevant, transferable governance implications for non-radiological imaging domains are noted to support broader applicability without overstating coverage. Many of the governance principles identified in this review, including bias mitigation, transparency, accountability, privacy protection, and lifecycle monitoring, are equally relevant to AI applications in pathology and cytology, although their implementation may require adaptation to modality-specific workflows, data characteristics, and diagnostic practices.

AI-driven imaging solutions increasingly demonstrate diagnostic performance comparable to expert clinicians in selected tasks and controlled evaluation settings [1,16,21,26–30]. Multimodal AI systems that integrate imaging data with electronic health records (EHRs), genomics, and clinical narratives signal a paradigm shift toward comprehensive AI-enabled diagnostic ecosystems. However, this technological progress introduces new sociotechnical complexities in relation to transparency, fairness, liability, and governance [27,31–49]. Opacity remains a central concern; despite high predictive performance, many models operate as “black boxes,” offering limited insight into decision pathways [14,50–68]. This

lack of transparency undermines clinician trust, complicates validation, and introduces medico-legal ambiguity [13,24,31,34,69–75].

Algorithmic bias further complicates the implementation, as unbalanced training data often perpetuate disparities between demographic groups [76–96]. These biases can lead to reduced diagnostic accuracy in minority populations, posing serious ethical and public trust challenges [19,23,44,48,69,70,97–100]. Thus, systematic governance frameworks for fairness auditing and subgroup validation are urgently needed [101–106]. Meanwhile, regulatory responses remain fragmented across jurisdictions. Despite promising advances such as the EU AI Act 2024 and the FDA’s SaMD framework, governance models often lag behind technical developments, especially adaptive, generative or multimodal AI systems [27,32,45,48,70,79,101,103,104,107]. Australia currently combines existing medical device regulation, professional guidance, and voluntary ethical principles; however, important governance challenges remain regarding adaptive AI systems, post-deployment oversight, fairness assurance, and accountability mechanisms [13,46,70,77,100,108–117]. Existing reviews have frequently focused on technical performance, algorithm development, or isolated ethical and regulatory challenges within diagnostic AI. Comparatively fewer studies have examined how ethical, governance, regulatory, and implementation considerations interact across the full AI lifecycle, from data acquisition and model development to deployment, monitoring, and post-market oversight. Furthermore, the rapid emergence of foundation models, generative AI systems, adaptive algorithms, and evolving regulatory frameworks has created new governance challenges that are not comprehensively captured in earlier reviews. This review addresses these gaps through an integrated synthesis of ethical, governance, regulatory, implementation, and trust-related considerations relevant to contemporary AI-enabled diagnostic imaging systems. To our knowledge, no previous review has comprehensively examined these interconnected dimensions across the full AI lifecycle while simultaneously considering foundation models, generative AI, multimodal systems, and emerging regulatory developments. In response, this systematic survey offers a comprehensive interdisciplinary synthesis of 156 sources (2018–2025), including peer-reviewed publications, regulatory documents, policy reports, and professional guidance materials, across four thematic domains: ethical considerations (52 sources), broader healthcare applications (57 sources), diagnostic imaging implementation (43 sources), and governance, policy, and education (4 sources) [12,71,96,118–125]. This categorisation is summarised in Table 1. The primary evidence corpus comprised sources published between 2018 and 2025. Earlier publications cited throughout the manuscript were used only to provide foundational, historical, conceptual, or regulatory context and were not included in the formal evidence synthesis.

Table 1. Distribution of Included Sources Across Thematic Domains and Relative Representation (2018–2025).

Thematic Domain	Sources (n)	Percentage (%)
Ethical Considerations	52	33.3
Broader Healthcare Applications	57	36.5
Diagnostic Imaging Implementation	43	27.6
Governance, Policy, and Education	4	2.6
Total	156	100.0

To examine the evolution of research in these domains, Table 2 presents the distribution of articles by year and thematic area from 2018 to 2025. This analysis reveals a steady increase in the level of scholarly attention to the ethical, governance, and policy dimensions of AI integration in medical diagnostics, in addition to continuing advances in broader

healthcare applications and diagnostic imaging. These trends highlight the growing recognition that governance frameworks must address not only technical performance but also transparency, fairness, and accountability to ensure responsible and equitable deployment of AI in clinical practice.

Table 2. Temporal Distribution of Included Sources Across Thematic Domains (2018–2025).

Year	Ethical Considerations	Broader Healthcare Applications	Diagnostic Imaging Implementation	Governance, Policy, Education	Integrated in This Study
2018	2	1	1	0	✓
2019	7	5	4	0	✓
2020	10	9	7	1	✓
2021	9	10	7	1	✓
2022	8	10	7	1	✓
2023	8	9	7	0	✓
2024	5	8	5	1	✓
2025	3	5	5	0	✓
Total	52	57	43	4	✓ (All)

1.2. Comparison with Previous Studies (2018–2025)

Although several reviews between 2018 and 2025 have examined AI in medical imaging, most have focused narrowly on technical performance, specific ethical concerns, or isolated governance aspects [16,23,24,103]. Few provide a comprehensive synthesis that integrates these dimensions between different modalities and jurisdictions [1,13,21]. This systematic survey addresses this gap by analysing 156 sources published between 2018 and 2025 and categorising them into ethical considerations, healthcare applications, diagnostic implementation, and governance frameworks. The trends captured in Table 2 reveal a growing shift toward ethical and policy-oriented concerns along with technical advances. Crucially, previous studies often overlooked emerging AI paradigms, such as foundation models, generative AI, multimodal fusion, and federated learning, which introduce new ethical and regulatory complexities [72,121,122]. Our analysis incorporates these developments, assessing their impact on key ethical and governance concerns. Although earlier reviews typically stop at general ethical discussions [31,32,70], we advance the field by offering concrete policy recommendations tailored to the Australian healthcare system, emphasising regulatory standardisation, adaptive oversight, fairness auditing, and stakeholder accountability.

1.3. The Role of AI in Medical Imaging

The integration of artificial intelligence (AI) into medical imaging represents one of the most consequential evolutions in diagnostic medicine [15,26,69,126–132]. Using large-scale annotated datasets and advanced computational architecture—such as convolutional neural networks (CNN), recurrent models, vision transformers, and generative frameworks—AI systems are capable of learning abstract, high-dimensional feature representations directly from raw pixel data [1,6,13,21,54]. These representations allow automated systems to perform complex image-based tasks with high fidelity, ranging from anatomical segmentation and lesion detection to disease classification, risk stratification, and outcome prediction [26,70,106,109–117,119,133,134].

1.3.1. Automated Disease Detection and Diagnosis

AI applications in medical imaging have achieved benchmark performance in numerous specialties. In oncology, AI models can detect pulmonary nodules, breast lesions,

and brain tumours with high sensitivity, often outperforming average radiologists in retrospective settings [1,12,13,15,16]. Similar success has been reported in cardiology and neurology, where deep learning systems identify ischaemic strokes, atherosclerotic plaque, and aneurysms from CT angiography and MRI scans [11,21,54,132,135]. In the context of infectious diseases, such as COVID-19, artificial intelligence has been deployed for the rapid detection of viral pneumonia from chest radiographs and lung CT, demonstrating clinical utility in the triage during high volumes of patients [11,13,98,100,107,108,135–138].

1.3.2. Radiomics and Prognostic Modelling

Beyond detection, AI enables radiomics, a domain focused on extracting and analysing quantitative imaging features that are imperceptible to the human eye [1,15,16,73,135]. These features, when combined with patient-specific clinical data, support precision diagnostics by identifying phenotypic signatures associated with treatment response, tumour progression, or survival outcomes. Predictive models trained on radiometric inputs are increasingly used in stratifying cancer patients, guiding the selection of personalised therapy, and forecasting the risk of recurrence in colorectal, breast, and prostate cancers [11,13,18,27,61,98,100,107,108,133–138].

1.3.3. Workflow Optimisation and Operational Efficiency

AI systems also contribute to efficiency in clinical workflows. Applications include automated case triage based on acuity, prioritisation of emergent conditions (e.g., intracranial haemorrhage), real-time quality control in imaging acquisition, and the generation of structured reports using natural language processing (NLP) [27,36,42,59,139]. AI-powered integration with Picture Archiving and Communication Systems (PACS) and Radiology Information Systems (RIS) can reduce reporting delays, improve workflow efficiency, and mitigate human error. These systems increasingly support multimodal imaging analysis and the fusion of images with clinical notes or omics data, enabling more holistic diagnostic insights [9,12,13,72,127].

1.3.4. Emerging Frontiers: Multimodal, Federated, and Generative Models

The next generation of AI imaging applications involves multimodal fusion: combining radiological data with genomics, pathology slides, or electronic health records to build richer diagnostic models [27,67,69,72,140]. Moreover, federated learning allows AI models to train in decentralised hospital datasets without compromising patient privacy, aligning with global data protection laws such as GDPR [25,27,44,120,141]. Generative adversarial networks (GANs) are also being explored to create synthetic training data sets, expand minority cases, and address bias by balancing underrepresented pathologies [72,98,121,122,125].

1.4. Ethical and Regulatory Challenges

The integration of artificial intelligence (AI) into medical imaging amplifies longstanding ethical concerns while introducing new regulatory complexities beyond traditional healthcare governance [11,12,26,70,79,101–106,142]. Across the reviewed literature, ethical concerns consistently clustered around five recurring domains: subgroup bias and inequitable performance, limited model interpretability in high-stakes clinical contexts, ambiguous accountability between clinicians and AI systems, privacy vulnerabilities in multi-institutional data ecosystems, and regulatory fragmentation across jurisdictions. While these concerns are widely acknowledged, operational mechanisms for systematic enforcement and lifecycle oversight remain comparatively underdeveloped [25,32,46,48,70,70,77,108,130,143].

As shown in Table 3, this taxonomy demonstrates that while technical solutions addressing bias, explainability, and privacy are frequently proposed, their validation often

remains retrospective rather than prospective. In contrast, governance mechanisms for post-deployment monitoring and adaptive model updates are comparatively underdeveloped. The distribution of validation maturity across lifecycle stages reveals a structural imbalance: innovation in algorithmic capability frequently outpaces institutional readiness for oversight and accountability.

Table 3. Solution Taxonomy of Technical and Governance Responses Across the AI Lifecycle, Including Validation Maturity and Implementation Limitations.

Solution Category	Representative Methods	Lifecycle Stage	Validation Level	Key Limitations
Bias Mitigation	Reweighting, adversarial debiasing, subgroup performance reporting	Development/Validation	Retrospective clinical testing common	Limited prospective deployment validation; dataset imbalance persists
Explainability (XAI)	SHAP, LIME, saliency maps, counterfactual explanations	Validation/Deployment	Simulated and retrospective studies predominant	Lack of standardised clinical interpretability metrics; uncertain impact on decision-making
Privacy-Preserving Learning	Federated learning, differential privacy, secure aggregation	Data acquisition/Training	Pilot and simulation studies	Trade-offs between privacy guarantees and diagnostic performance
Post-Deployment Monitoring	Performance dashboards, drift detection algorithms, audit logging	Deployment/Monitoring	Emerging implementation evidence	Sparse regulatory standardisation; unclear escalation protocols
Regulatory Governance Tools	SaMD approval pathways, post-market surveillance, PCCP frameworks	Deployment/Lifecycle update	Policy guidance evolving	Limited integration with adaptive and continuously learning systems

Table 3 demonstrates that while numerous technical and governance interventions have been proposed across the AI lifecycle, many remain supported primarily by retrospective, pilot, simulation, or policy-level evidence, highlighting the need for stronger prospective clinical validation and post-deployment evaluation.

1.4.1. Algorithmic Bias and Fairness

Bias often arises from unrepresentative data, skewed sampling, and feedback loops that reinforce disparities, disproportionately affecting marginalised groups [6,23,69,98,100,106,125,137]. Fairness requires inclusive datasets, continuous bias audits, demographic performance reporting, and active stakeholder engagement throughout development [19,24,26,74,143,144].

1.4.2. Transparency and Explainability

Deep learning models often lack interpretability, compromising clinician trust, informed consent, and accountability [34,50,73,117,124,126,127]. Clinicians need not only accurate predictions but clear reasoning, especially in critical cases such as oncology and surgical planning [24,40,72,77,144]. Tools such as saliency maps, LIME, SHAP, and counterfactuals show promise but remain inconsistent and lack clinical validation standards [14,25,72,73,119,126,134].

1.4.3. Accountability and Liability

Current legal frameworks struggle to assign responsibility when AI errors cause harm [31,33,79,100,101,145]. This is particularly urgent for autonomous or semi-autonomous systems, such as triage algorithms [12,51,103,114]. Clear accountability frameworks are needed, including audit trails, consent processes that recognise algorithmic uncertainty, and third-party validation in clinical settings [17,32,107,108].

1.4.4. Regulatory Fragmentation and Global Inconsistencies

Disparities between regulatory bodies such as the FDA, EMA, and Australia's TGA complicate the oversight of AI-enabled medical devices [45,70,103,104,107,108]. Jurisdictions differ particularly in the regulation of adaptive models that evolve after deployment [35,46,48,70,146]. Meanwhile, ethics frameworks often do not keep pace with technology, weakening transparency and accountability [23,49,50,73,86,100].

These challenges highlight the need for globally coordinated and interdisciplinary governance. Equitable data governance, participatory policy design, and ethically embedded technical development are crucial [71,96,118–125]. As conceptually illustrated in Figure 1, the reviewed literature suggests that ethical concerns surrounding AI in medical imaging became prominent earlier than the development of formal governance frameworks, highlighting a persistent gap between ethical expectations and regulatory readiness.

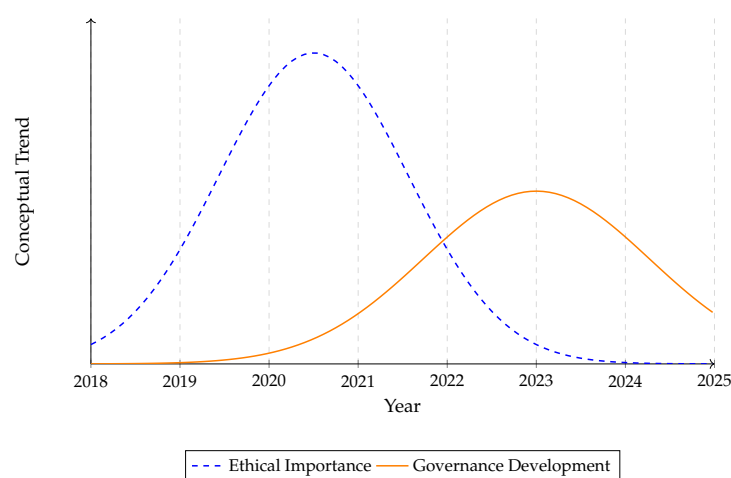


Figure 1. Conceptual illustration of the evolving relationship between ethical concerns and governance development in AI-enabled medical imaging between 2018 and 2025, based on thematic patterns identified in the reviewed literature.

1.5. Comparative Analysis of Global Governance Frameworks

The governance of AI in healthcare is evolving unevenly across jurisdictions, shaped by regional policy priorities, institutional capacity, and historical regulatory paradigms. Although some regions are adopting proactive, risk-based legislation (e.g., the European Union AI Act 2024), others continue to operate through legacy medical device pathways or advisory ethical frameworks. This regulatory fragmentation creates a significant asymmetry in legal enforceability, ethical accountability, and innovation scalability [32,70,70,101,104,107].

In all jurisdictions, common goals have emerged such as human oversight, explainability, transparency, and patient safety. However, the degree to which these goals are legally binding, operationalised, and aligned with clinical realities remains inconsistent. Emerging issues such as adaptive algorithmic updates, synthetic data training, and cross-border compliance raise additional complexity for regulators worldwide [6,27,36,45,77,127]. A comparative overview of AI governance strategies related to healthcare from ten representative frameworks is presented in a later section of this review. This comparison incorporates both binding legislation and normative ethical guidelines. This comparative lens underscores the need for transnational coordination, agile co-regulation models, and standards for continuous algorithmic validation—particularly as AI systems transition from static tools to dynamic, learning entities deployed at scale [15,26,69,79,100,101,107,108,126–132].

Artificial intelligence (AI) is rapidly transforming medical imaging by enhancing diagnostic precision, reducing clinician burden, and optimising radiological workflows [11,12]. During the past decade, AI techniques, particularly deep learning (DL), convolutional

neural networks (CNNs), and natural language processing (NLP), have enabled automated image classification, lesion detection, radiomics, and clinical report generation [70]. Although these capabilities have achieved state-of-the-art performance in various tasks, they also raise urgent ethical and regulatory concerns [13,34,69,101,142]. In the context of the integration of AI into radiology and diagnostic decision making, it is critical to establish governance mechanisms that address societal concerns and clinical safety. Current debates around artificial intelligence in healthcare protocols increasingly emphasise not only technical precision but also the development of systems that are explicable, equitable, and ethically grounded [11,46,97].

1.6. Key Advances in AI-Driven Medical Imaging

The application of AI in medical imaging has evolved from discrete detection tasks to complete, end-to-end diagnostic support systems. These systems combine computer vision techniques with clinical reasoning tools, underpinned by convolutional neural networks (CNNs), transformers, radiomics, and large language models (LLMs) [11–13,15,16,71,90].

- **Automated Image Interpretation:** AI models now reliably identify tumours, fractures, nodules, and haemorrhages from CT, MRI, PET and X-ray scans, performing on the same level as experienced radiologists in many settings [1,40,105].
- **Radiomics and Predictive Modelling:** Advanced radiomic workflows extract latent image features to stratify patient risk, predict recurrence, and assess therapeutic responsiveness in oncology and neurology [16,70,73].
- **Clinical Workflow Optimisation:** AI augments PACS/RIS systems by prioritising emerging findings, automating measurements, and populating structured radiology reports, thus improving operational efficiency [27,36,139].
- **Multimodal Integration via Foundation Models:** Emerging LLM-powered platforms synthesise textual and imaging data, allowing queries in natural language, diagnostic summarisation, and report generation [71,90,140].

AI models serve as intermediaries between raw image data and downstream diagnostic decision making. Feedback loops ensure continuous learning and calibration of model predictions based on the input of the radiologist, enhancing robustness and reliability [15,25,40,127].

1.7. Visualising AI's Role in Medical Imaging

To capture the intricacies of AI integration within diagnostic imaging, Figure 2 presents a comprehensive modular pipeline that represents the flow from raw image acquisition to clinical decision support. This enhanced schematic emphasises not only the technical progression of the data, but also the integration of essential components such as explainability, privacy, and continuous learning, reflecting current research trends and regulatory priorities. The workflow begins with *Image Acquisition*, where imaging modalities such as CT, MRI, PET, or X-ray produce diagnostic scans. These raw images are then passed to the *Preprocessing* layer, which applies normalisation, denoising, and contrast enhancement to ensure input standardisation and reduce artefacts. This step may also include anatomical labelling and harmonisation, especially in multicentre deployments [71,74,96,97,99,118–125,139,143,144,146–149].

Following preprocessing, data enter the AI Inference Engine, which may consist of convolutional neural networks (CNNs), transformers, or large multimodal language models (LLMs). These models perform critical computational tasks such as tumour segmentation, anomaly detection, risk stratification, and report generation. In parallel, an integrated explainability module operates, employing techniques such as SHAP, LIME, saliency maps, and counterfactuals to generate human-understandable justifications for model outputs. This is essential for transparency, regulatory compliance, and the promotion of patient trust [27,41–49]. A dedicated Privacy Protection Layer protects sensitive health data through

methods such as federated learning, differential privacy, and homomorphic encryption, ensuring compliance with GDPR, HIPAA, and related standards without degrading AI performance. The output of the inference engine is then delivered to a Diagnosis Output unit, which produces structured results (e.g., classifications, segmentations, heatmaps) for clinical interpretation. These results are seamlessly integrated into hospital systems such as PACS (Picture Archiving and Communication Systems) or RIS (Radiology Information Systems) through a Clinical Decision Support module, allowing real-time triage, prioritisation of urgent findings, and partially or fully automated reporting [70,109–117].

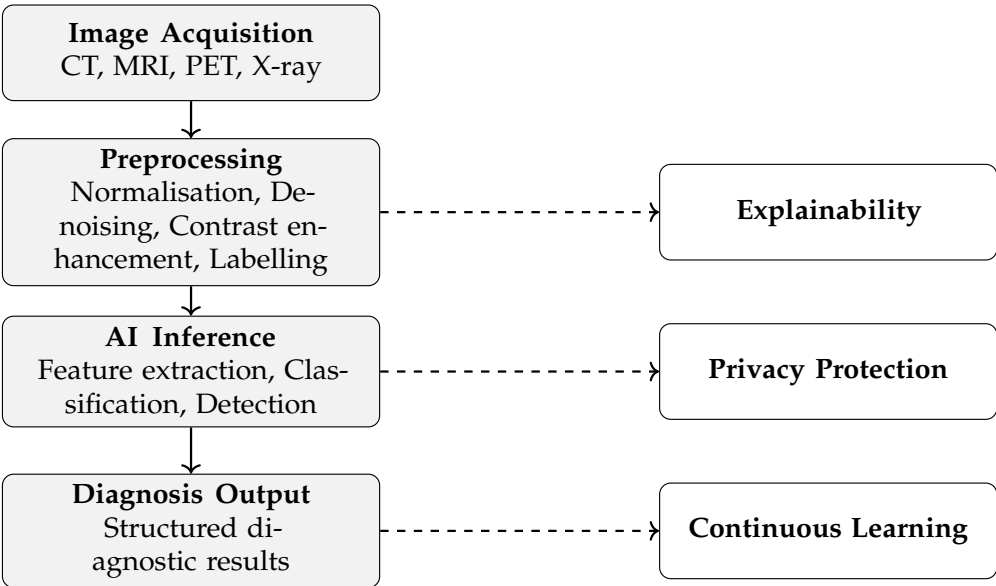


Figure 2. Modular AI pipeline for medical imaging illustrating the progression from image acquisition to diagnostic output and highlighting key governance considerations, including explainability, privacy protection, and continuous learning.

A crucial component of this pipeline is the human-in-the-loop feedback system, in which radiologists review AI-generated outputs, validate results, and provide corrective feedback. This continuous feedback loop supports model retraining, mitigates performance drift, and promotes ongoing learning aligned with evolving clinical practice. By incorporating explainability, privacy safeguards, and feedback-driven iteration, the system examines not only technical robustness, but also ethical, legal, and social requirements essential for reliable AI deployment in healthcare. Figure 2 shows a generalised workflow that illustrates the full pipeline: from raw medical image acquisition (MRI, CT, X-ray) through preprocessing and feature extraction to AI-driven analysis and diagnostic output. The modular structure reflects the interdependent stages of AI-assisted diagnostic radiology, underscoring the need for data integrity, transparency, and accountability at each step [76–85]. Table 4 summarises this functional architecture, detailing how data flow across acquisition, preprocessing, inference, and decision support while embedding safeguards for explainability, privacy, and continuous learning. The deployment of AI in medical imaging, while technically transformative, demands a robust and multilayered governance framework to ensure ethical sustainability and clinical safety. Governance must extend beyond voluntary codes or retrospective audits, establishing enforceable mechanisms guiding the entire AI lifecycle, from data acquisition and model development to deployment, ongoing monitoring, and decommissioning. This framework must remain proactive and adaptive, ready to address the emerging challenges of rapidly evolving technologies such as LLM and generative AI [72,73,86–95,133,134,140,141,150–153].

Table 4 illustrates that ethical and governance considerations emerge throughout the AI-enabled diagnostic workflow. These considerations extend from data acquisition and preprocessing to model explainability, clinical deployment, and post-deployment learning processes.

As illustrated in Figure 2, governance considerations are not confined to a single stage of the workflow but must be embedded throughout the AI lifecycle, spanning data acquisition, model development, deployment, and post-deployment monitoring.

Table 4. Key Functional Modules and Associated Technologies in AI-Enabled Medical Imaging.

Module	Core Functions	Illustrative Technologies
Image Acquisition	Captures diagnostic scans from various modalities	CT, MRI, PET, X-ray
Preprocessing	Enhances image quality, reduces artifacts, standardises inputs	Denosing, normalisation, contrast harmonisation
AI Inference Engine	Analyses patterns, detects abnormalities, generates predictions	CNNs, Transformers, LLMs, radiomics
Explainability Module	Improves transparency and model interpretability	SHAP, LIME, saliency maps, counterfactuals
Privacy Layer	Secures sensitive data and protects patient identity	Federated learning, homomorphic encryption
Clinical Decision Support	Delivers actionable insights and integrates into workflows	PACS/RIS integration, triage, automated reporting
Feedback Loop	Enables continuous learning from radiologist input	Human-in-the-loop validation and retraining

1.8. Derivation of the Conceptual Governance Framework from the Surveyed Literature

The five-pillar conceptual ethical governance framework proposed in this survey is derived from the lifecycle-centric synthesis of ethical risks and governance responses identified across the reviewed literature. As the framework emerged from thematic synthesis of the included evidence rather than empirical implementation or prospective clinical validation, it is presented as a conceptual governance framework intended to organise and operationalise governance requirements identified across the evidence base. Drawing on the taxonomy presented in Table 5, recurring governance gaps were traced across data curation, model development, validation, clinical deployment, and post-deployment monitoring stages. The resulting pillars reflect intervention points consistently emphasised across studies, rather than abstract normative principles introduced independently of the survey findings.

The proposed conceptual framework does not seek to introduce entirely new ethical principles. Rather, it synthesises and operationalises governance requirements repeatedly identified across the reviewed literature and aligns them with established international guidance. Existing frameworks, including the WHO Guidance on Ethics and Governance of Artificial Intelligence for Health, OECD AI Principles, FDA Good Machine Learning Practice guidance, the European Union AI Act, and Australian TGA recommendations, provide important ethical principles, regulatory obligations, and risk-management expectations. However, these frameworks are predominantly presented as principle-based, regulatory, or risk-classification approaches and provide limited guidance on how governance activities should be systematically integrated across the full AI lifecycle.

The principal contribution of the proposed conceptual framework lies in its lifecycle-centric organisation of governance requirements. By mapping ethical risks, governance responses, accountability mechanisms, and trust indicators to specific stages of the AI lifecycle, the framework translates broad ethical principles into operational governance activities. This structure enables continuous oversight across data acquisition, model development,

validation, clinical deployment, post-deployment monitoring, and adaptive system updating. In doing so, the framework provides a practical mechanism for identifying governance gaps, assigning accountability, and supporting trust calibration throughout the lifecycle of AI-enabled diagnostic systems.

Table 5. Lifecycle-Centric Taxonomy of Ethical and Governance Challenges, Governance Responses, and Trust Indicators in AI-Enabled Medical Imaging.

AI Lifecycle Stage	Dominant Ethical Risks	Commonly Proposed Governance Responses	Key Limitations Identified in the Literature	Illustrative Trust Indicators
Data Collection and Curation	Dataset bias, lack of representativeness, secondary data use without consent, privacy leakage	Data governance policies, consent frameworks, bias audits, de-identification standards	Limited patient involvement, weak enforcement mechanisms, inadequate coverage of minority populations	Dataset representativeness, demographic coverage, patient consent participation rates
Model Development and Training	Algorithmic bias amplification, opacity of decision logic, reproducibility concerns	Model documentation, explainability tools, interdisciplinary review committees	Limited clinical validation of explainability methods, lack of standardisation across institutions	Calibration error, confidence reliability, subgroup performance consistency
Validation and Regulatory Approval	Inadequate external validation, misalignment with clinical workflows, unclear accountability	Regulatory certification processes, clinical trials, performance benchmarking	Static approval models that fail to address adaptive or continuously learning systems	Radiologist–AI concordance, external validation stability across datasets
Clinical Deployment and Use	Automation bias, over-reliance on AI outputs, unclear clinician responsibility	Human-in-the-loop requirements, professional guidelines, role delineation protocols	Ambiguous liability allocation, variable uptake across clinical contexts	Clinician adoption rate, override frequency, real-world reliance patterns
Post-Deployment Monitoring and Adaptation	Model drift, performance degradation, governance gaps for system updates	Post-market surveillance, audit trails, continuous monitoring requirements	Limited regulatory guidance for adaptive systems, fragmented oversight responsibilities	Patient acceptance rates, incident-triggered review, longitudinal trust feedback trends

Table 5 demonstrates that ethical risks, governance requirements, and trust indicators vary across the AI lifecycle, reinforcing the need for stage-specific oversight mechanisms rather than a single governance approach applied uniformly across all phases of AI development and deployment.

Table 6 maps each governance pillar to the evidence signals identified across the reviewed literature (n = 156), their relative prevalence, and the principal implementation gaps reported in the evidence base.

Table 6. Mapping of Governance Pillars to Evidence Signals in the Reviewed Literature (n = 156).

Governance Pillar	Evidence Signals in Literature	Prevalence	Observed Gaps
Equity and Bias Mitigation	Subgroup performance reporting, dataset representativeness analysis, fairness metrics	High	Limited prospective validation; underrepresentation of minority populations
Explainability and Transparency	Use of SHAP, LIME, saliency maps, counterfactual explanations	Moderate	Limited clinical validation and standardisation of interpretability metrics
Accountability and Oversight	Human-in-the-loop review, audit trails, regulatory documentation, clinical governance committees	Moderate	Unclear delineation of clinician vs. AI responsibility; limited post-market accountability mechanisms
Privacy-Preserving Infrastructure	Federated learning, differential privacy, secure multiparty computation	Emerging	Limited deployment evidence; trade-offs with model performance insufficiently evaluated
Adaptive Regulatory Alignment	Lifecycle monitoring proposals, post-deployment surveillance, change control protocols	Low–Moderate	Sparse empirical guidance for managing model drift and continual learning systems

1. **Bias Mitigation and Equity Auditing**
 - Address algorithmic disparities caused by skewed or incomplete training data.
 - Mitigate diagnostic bias between race, gender, and age groups [69,99,100].
 - Insist on equity audits, demographic benchmarking, and subgroup performance reporting.
2. **Explainability and Transparency**
 - Combat the opacity of black-box models (e.g., deep neural networks).
 - Integrate explainable AI (XAI) tools saliency maps, SHAP values, counterfactuals to clarify predictions [26,34,73].
 - Make interpretability accessible to both clinicians and patients.
3. **Interdisciplinary and Participatory Oversight**
 - Involve a broad stakeholder base: Clinicians, ethicists, patients, legal experts, and developers.
 - Promote ethical co-design to improve foresight and system resilience [46,86,97].
4. **Privacy-Preserving Data Infrastructure**
 - Embed privacy by design using secure techniques: federated learning, differential privacy, and homomorphic encryption [27,44,120].
 - Ensure compliance with data protection laws (e.g., GDPR, HIPAA) and minimise re-identification risk.
5. **Regulatory and Policy Alignment**
 - Harmonise ethical AI practices with global frameworks: EU AI Act 2024, FDA SaMD guidelines, WHO recommendations [70,101,107].
 - In Australia, bridge policy gaps by advancing national regulatory standards.

The prevalence classifications reflect thematic frequency across the 156 reviewed studies. Ethical concerns related to bias and fairness were widely discussed, whereas structured governance mechanisms for post-deployment monitoring and adaptive systems were comparatively underdeveloped. This distribution highlights a recurring pattern in the literature: normative recognition of risks often exceeds operationalised governance implementation.

1.9. Real-World Deployment and Governance Challenges in Clinical Practice

While much of the surveyed literature focuses on anticipated ethical risks, several real-world deployment experiences illustrate how governance gaps materialise in clinical settings. These cases highlight the importance of lifecycle-aware oversight and reinforce the practical relevance of the proposed governance framework.

1.9.1. Case 1: Performance Degradation and Post-Market Oversight

Several AI-enabled radiology tools have demonstrated reduced diagnostic performance when deployed outside their original training environments, particularly when applied across different patient populations or imaging devices. Such cases expose governance gaps at the post-deployment monitoring stage, where model drift and context shift are insufficiently addressed by static approval mechanisms. These experiences underscore the need for continuous performance auditing and adaptive oversight, directly motivating the framework's emphasis on lifecycle-aware governance and accountability structures.

1.9.2. Case 2: Automation Bias and Clinical Responsibility

Reports of clinicians over-relying on AI-generated outputs in diagnostic workflows have raised concerns regarding automation bias and unclear responsibility boundaries. In these instances, AI recommendations were followed despite conflicting clinical judgement,

revealing weaknesses in human-in-the-loop governance during the clinical deployment stage. The literature consistently identifies the absence of clear role delineation and accountability protocols as contributing factors, reinforcing the framework's focus on clinician oversight, responsibility allocation, and institutional governance safeguards.

1.9.3. Case 3: Data Governance and Secondary Use Risks

Real-world controversies surrounding the secondary use of medical imaging data for AI development without meaningful patient awareness or consent have highlighted persistent data governance failures. These cases, often emerging during data curation and post-deployment data reuse stages, demonstrate how privacy risks extend beyond initial model development. Such examples support the framework's prioritisation of consent-aware data stewardship and transparent governance mechanisms that extend across the full AI lifecycle.

- **Opacity and the Epistemic Black Box:** Clinicians often report difficulty in interpreting model predictions or assessing their reliability in atypical cases [13,34,71,72]. This epistemic opacity undermines informed consent, impedes error investigation, and complicates medico-legal accountability.
- **Algorithmic Fairness and Diagnostic Disparities:** The lack of fairness metrics in performance reporting perpetuates these blind spots. Bias can also arise from equipment heterogeneity, annotation subjectivity, and socioeconomic data gaps [69,97,99].
- **Unclear Legal and Professional Accountability:** Should a misclassification be attributed to the radiologist, the AI developer, or the deploying hospital? Current legal frameworks remain inadequate to address these distributed responsibilities [2,19,31,33]. Without clear attribution mechanisms, accountability becomes diffused, increasing the risk of untraceable harm.
- **Privacy, Consent, and Data Sovereignty:** The use of data in AI training and inference requires stringent privacy safeguards. However, many systems are built on retrospective datasets collected without explicit consent for AI use [27,44,101,120]. Furthermore, cloud-based services used to train or deploy AI models can transfer data across jurisdictions, raising concerns about data sovereignty and legal exposure.
- **Regulatory Ambiguity and Lifecycle Oversight:** Regulatory frameworks have yet to fully address the continuous evolution of AI models, particularly those that update after deployment through online learning or user feedback. This adaptive nature introduces new risks, such as concept drift and unintended behaviour changes that escape initial validation [11,46,70,107,108,142]. The absence of standardised lifecycle governance that covers monitoring, revalidation, and retirement leaves a critical oversight gap.
- **Risk of Clinical Deskilling and Over-Reliance:** Studies have shown that frequent deferral to automated recommendations can diminish diagnostic intuition and reduce the likelihood of independent verification, especially under time constraints [41,42,139]. This phenomenon raises long-term questions about professional competence and patient safety.
- **Socio-ethical Disconnects and Lack of Public Trust:** The socioethical disconnect contributes to mistrust, particularly when patients do not know that their care is being influenced by algorithms [44,50,58]. Without transparent communication and participatory design, AI risks deepening existing fault lines in the patient–clinician relationship.

These ethical challenges are not static; they evolve as AI systems become more sophisticated and embedded in clinical infrastructure. Addressing them requires not only technical interventions, but also institutional readiness, legal reform, and cultural shifts

toward transparency and inclusion. Governance, therefore, must operate at the intersection of ethics, policy, and real-world clinical practice. Figure 3 presents a conceptual alignment between the core governance mechanisms and the main ethical challenges associated with AI-enabled medical imaging. It consists of two parallel vertical structures: the left side highlights five foundational pillars of ethical governance, and the right side lists the major ethical challenges commonly encountered in AI deployments. Each directional arrow represents a targeted intervention by which a governance component is intended to mitigate or directly address a corresponding risk or limitation.

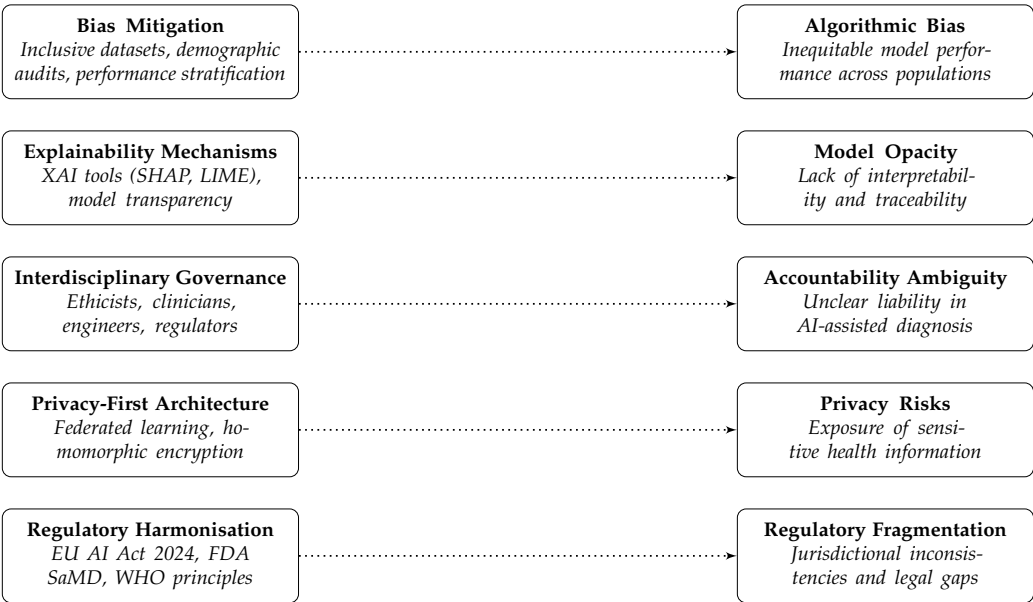


Figure 3. Conceptual alignment between ethical governance mechanisms and the principal ethical and operational challenges encountered in AI-enabled medical imaging.

Figure 3 illustrates that effective governance requires targeted interventions that directly address specific ethical and operational risks, rather than relying on generic governance controls applied uniformly across the AI lifecycle.

- **Bias Mitigation (Algorithmic Bias).** Bias mitigation strategies involve curating demographically inclusive datasets and applying subgroup-level performance tests to prevent discriminatory results. This governance pillar responds to algorithmic bias, a challenge stemming from unbalanced or non-representative training data, which may lead to diagnostic inaccuracies in under-represented groups [69,99,121–125].
- **Explainability (Opacity of AI Systems).** The use of model-agnostic interpretability techniques such as SHAP (Shapley Additive Explanations), LIME (Local Interpretable Model-Agnostic Explanations), and saliency maps aims to make opaque AI systems clearer to clinicians. This evaluates the lack of interpretability associated with black-box models, particularly deep neural networks, and supports informed clinical use and regulatory scrutiny [26,73,97,99,139,144,147,148].
- **Interdisciplinary Oversight (Accountability Ambiguity).** Establishing shared governance across clinical, ethical, legal and technical domains enables clarity around responsibility and liability in AI-enabled diagnostic pathways. This is critical to solving the accountability vacuum that arises when AI errors occur, especially when models operate autonomously or in critical contexts [31,79].
- **Privacy-first infrastructure (data privacy violations).** Federated learning, differential privacy, and homomorphic encryption are part of a privacy-first architectural approach that protects sensitive patient data during AI training and deployment. These safe-

guards directly respond to concerns about unauthorised access and reidentification risks in distributed or cloud-based imaging environments [27,44].

- Policy alignment (Regulatory fragmentation). Policy harmonisation with frameworks such as the EU AI Act 2024, the US FDA's SaMD Action Plan and WHO's AI ethics guidance is vital to bridge inconsistencies between jurisdictions in regulation. This ensures consistent lifecycle oversight, improves interoperability, and fosters transnational trust in AI systems deployed in various health systems [70,107]. The figure underscores how a proactive and modular governance strategy can preemptively mitigate downstream risks, creating the ethical infrastructure necessary for a responsible, scalable, and clinically safe deployment of AI in medical imaging.

Taken together, these ethical challenges do not arise in isolation but emerge across the full AI lifecycle in medical imaging, from data acquisition and model development to deployment and post-market monitoring. Issues frequently co-occur, reinforcing one another and complicating clinical governance. Rather than treating these concerns as discrete risks, the literature increasingly frames them as interdependent sociotechnical failures requiring coordinated regulatory, institutional, and technical responses.

2. Review Plan and Conduct

This systematic survey of the literature was conducted and reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews (PRISMA-ScR) guidelines where applicable to interdisciplinary governance-focused evidence synthesis.

Although this study employed a systematic search strategy, predefined eligibility criteria, and structured evidence synthesis procedures, PRISMA-ScR reporting guidance was selected instead of the standard PRISMA framework because the primary objective of this review was to map and synthesise a broad and heterogeneous body of evidence rather than evaluate the effectiveness of a specific intervention or undertake quantitative outcome comparison. The review incorporated diverse evidence sources, including peer-reviewed studies, regulatory documents, policy reports, and professional guidance materials spanning clinical, technical, ethical, and governance domains. Given this interdisciplinary scope and the emphasis on identifying, categorising, and synthesising ethical and governance challenges across the AI lifecycle, PRISMA-ScR was considered more appropriate than standard PRISMA, which is primarily designed for systematic reviews focused on intervention effectiveness and quantitative evidence synthesis. Accordingly, this study is presented as a systematic survey of the literature informed by PRISMA-ScR reporting principles.

A completed PRISMA-ScR checklist, including corresponding manuscript page references, has been provided as Supplementary File S1 to support reporting transparency and facilitate assessment of compliance with PRISMA-ScR reporting recommendations. Full database-specific search strategies and corresponding search dates have been provided in Supplementary File S2 to enhance methodological transparency and reproducibility.

While systematic search, screening, eligibility assessment, and structured synthesis procedures were employed, the review was designed to identify, classify, and synthesise heterogeneous evidence relating to ethical, governance, regulatory, technical, and implementation challenges in AI-enabled diagnostic imaging. Consistent with the objectives of a systematic survey, the emphasis was placed on evidence mapping, thematic synthesis, and governance analysis across diverse evidence types rather than quantitative outcome comparison or intervention-effectiveness assessment.

The methodology was designed to support interdisciplinary thematic synthesis and lifecycle-oriented governance analysis across scholarly, clinical, ethical, and regulatory literature relating to AI-enabled diagnostic imaging. This research design was selected

because the study sought to systematically identify, classify, and synthesise heterogeneous evidence spanning technical, clinical, ethical, governance, and regulatory domains. A PRISMA-ScR-informed systematic survey was considered the most appropriate approach for addressing the research questions because it enables broad evidence mapping, thematic synthesis, and lifecycle-oriented governance analysis across diverse evidence types that would not be suitable for quantitative review or meta-analysis.

To ensure a robust and thematically comprehensive methodological foundation for this review, this section details the systematic approach used in the identification, selection, classification, and synthesis of the relevant literature. Given the interdisciplinary nature of AI in medical imaging, spanning computer science, clinical medicine, ethics, and regulatory science, the review protocol was designed to accommodate heterogeneity across sources while preserving reproducibility and conceptual coherence.

The review was designed to capture not only technical model development but also scholarship addressing AI integration within digital health systems, including regulatory policy, health informatics governance, clinical workflow redesign, and institutional accountability structures.

2.1. Objectives and Research Questions

The overarching objective of this review is to investigate the ethical and governance implications of AI integration in diagnostic medical imaging. The study is guided by the following research questions:

- **RQ1:** What are the dominant ethical challenges associated with the clinical deployment of AI in medical imaging?
- **RQ2:** How do national and international governance frameworks address issues such as accountability, transparency, and patient rights in AI-driven diagnostics?
- **RQ3:** What practical recommendations can be synthesised to support the integration of AI based on ethical principles and policy within the Australian healthcare system and beyond?

These questions form the backbone of the literature selection, coding taxonomy, and thematic analysis presented in the subsequent sections.

2.2. Search Strategy and Data Sources

An integrative multiphase search strategy was implemented to capture the interdisciplinary literature at the intersection of artificial intelligence, medical imaging, ethics, and health policy. The search covered publications from January 2018 to April 2025 and was designed to capture both foundational and contemporary contributions relevant to the evolving AI landscape. Recognising the rapid advancement of AI technologies and regulatory frameworks, this timeframe provided a balance between historical context and current relevance. All database searches were conducted and finalised in April 2025, after which the retrieved records were exported for screening, eligibility assessment, and thematic synthesis.

The search was carried out in three iterative phases to maximise coverage and minimise omission bias.

1. **Systematic Database Search:** Core academic repository was queried using controlled vocabulary and Boolean keyword combinations. The databases included PubMed/MEDLINE, ACM Digital Library, IEEE Xplore, SpringerLink, Elsevier ScienceDirect, and Google Scholar. PubMed/MEDLINE was included to ensure comprehensive coverage of biomedical, clinical, and healthcare-focused literature, while the remaining databases captured technical, interdisciplinary, and policy-oriented research. Keywords were selected to reflect both technical and normative domains (e.g.,

‘AI in radiology’, ‘algorithmic bias’, ‘SaMD regulation’, ‘explainability in diagnostics’), including trust-related terms such as ‘trust’, ‘adoption’, ‘acceptance’, ‘calibration’, and ‘clinician–AI concordance’.

2. **Snowball Sampling:** To ensure the inclusion of high impact and often cited studies, backward snowballing (mining reference lists of key papers) and forward snowballing (identifying cited works) were used. This phase enriched the dataset with both seminal and emergent literature not captured through standard indexing.
3. **Gray Literature and Policy Integration:** In recognition of the growing influence of regulatory bodies and think tanks in shaping AI governance, authoritative non-academic sources were incorporated. These included white papers and guidance documents from the World Health Organisation (WHO), the OECD, the European Commission, the US Food and Drug Administration (FDA), and the Therapeutic Goods Administration of Australia (TGA), among others.

This multisource, temporally inclusive search protocol ensured that the resulting literature corpus was comprehensive and representative of the diverse scholarly, regulatory, and clinical perspectives surrounding AI in medical imaging.

2.3. Inclusion and Exclusion Criteria

To maintain thematic integrity and ensure academic rigor, we applied clear eligibility parameters during the screening process. These criteria were designed to include interdisciplinary perspectives in ethics, governance, and AI applications in healthcare. The primary evidence corpus comprised sources published between January 2018 and April 2025. Earlier publications cited throughout the manuscript were used only to provide foundational, historical, conceptual, or regulatory context and were not included in the formal evidence synthesis.

- **Inclusion Criteria:** Peer-reviewed journal articles, regulatory documents, white papers, and institutional reports published between January 2018 and April 2025. The included works addressed the ethical, legal, governance, or sociotechnical dimensions of AI in medical diagnostics, with specific relevance to imaging modalities.
- **Exclusion Criteria:** Technical AI research lacking ethical, clinical, or policy considerations; publications not related to healthcare domains (e.g., industrial or military AI); and non-English-language sources. English-language publications were included to ensure consistent screening, interpretation, and thematic synthesis across a large interdisciplinary body of literature spanning technical, clinical, ethical, legal, and policy domains. Given resource and translation constraints, non-English publications were excluded. While this may have resulted in the omission of some relevant studies, English-language sources were considered sufficient to capture major international developments in AI ethics, governance, and diagnostic imaging.

2.4. Keyword Matrix and Search Query Design

The search strategy was guided by a structured matrix of keyword clusters aligned with the conceptual pillars of the study. This matrix ensured comprehensive semantic coverage and supported consistent query formulation across diverse databases.

- **AI Application Domains:** “medical imaging”, “radiology AI”, “diagnostic automation”, “deep learning in healthcare”, “computer-aided diagnosis”
- **Ethical Constructs:** “algorithmic bias”, “AI fairness”, “explainability”, “accountability in AI”, “clinical transparency”, “informed consent”
- **Governance and Legal Constructs:** “AI regulation in healthcare”, “FDA AI policy”, “EU AI Act 2024 healthcare”, “SaMD compliance”, “data protection law”, “health AI governance”

Database-specific syntax (e.g., proximity operators, Boolean logic) was used to tailor query execution. Synonyms and regional variations (e.g., “ethics” and “bioethics”, “regulation” and “compliance”) were included to ensure cross-disciplinary relevance.

2.5. Screening and Classification Protocol

All retrieved sources were imported into a central reference manager and screened at the level of title, abstract, and full text. Eligibility was assessed by two independent reviewers. Disagreements regarding study eligibility or thematic classification were resolved through discussion and reviewer consensus. During title and abstract screening, records were assessed against predefined eligibility criteria to determine their relevance to AI-enabled diagnostic imaging and associated ethical, governance, regulatory, clinical, or implementation considerations. Potentially eligible records subsequently underwent full-text review. Sources were included when they provided substantive discussion of ethical, legal, governance, regulatory, technical, clinical, or sociotechnical issues relevant to the development, deployment, oversight, or responsible implementation of AI in diagnostic imaging. The selected articles were coded using a tripartite classification system:

- a. **Ethics-Focused:** Literature emphasising bias mitigation, transparency, patient autonomy, or clinician-patient trust.
- b. **Governance-Orientated:** Works discussing regulatory instruments, legal liability, institutional policy, or compliance frameworks.
- c. **Clinical Utility:** Studies examining the impact of AI in the real world on diagnostic performance, efficiency, or clinical workflows.

This taxonomy facilitated cross-sectional synthesis and enabled the triangulation of ethical principles with policy and practice dimensions.

2.6. Thematic Domain Classification

Following full-text review, all included sources were categorised according to their primary thematic focus. An inductive thematic classification process was used to organise the literature into four overarching domains: (1) Ethical Considerations, (2) Broader Healthcare Applications, (3) Diagnostic Imaging Implementation, and (4) Governance, Policy, and Education.

Studies examining fairness, transparency, accountability, privacy, trust, informed consent, and related ethical issues were classified under Ethical Considerations. Publications exploring AI applications across broader healthcare contexts beyond diagnostic imaging were categorised as Broader Healthcare Applications. Studies focusing on the development, validation, deployment, or clinical integration of AI systems within diagnostic imaging workflows were grouped under Diagnostic Imaging Implementation. Sources primarily addressing regulatory frameworks, governance mechanisms, policy development, professional guidance, workforce preparedness, or educational considerations were assigned to Governance, Policy, and Education.

Where publications addressed multiple thematic areas, classification was based on the dominant focus identified during full-text review. Any uncertainties regarding thematic assignment were discussed among the review team and resolved through consensus. Thematic classification was undertaken through an iterative coding and synthesis process. Recurring concepts identified across the included sources were initially coded and subsequently grouped into higher-order thematic categories. These concepts included fairness and bias, explainability and transparency, accountability and oversight, privacy and data governance, regulatory alignment, clinical implementation, stakeholder trust, and post-deployment monitoring. The resulting thematic structure informed both the synthesis of governance challenges presented throughout the review and the development of the proposed conceptual governance framework.

2.7. Study Selection Summary

The database search initially yielded 642 records. After removal of duplicates (98), 544 titles and abstracts were screened. Full-text assessment was conducted for 213 articles, resulting in 156 studies included in the final synthesis.

Exclusions at the full-text stage were primarily due to: (i) studies focused solely on algorithmic performance without ethical or governance analysis, (ii) AI applications outside diagnostic imaging, (iii) non-peer-reviewed sources, and (iv) insufficient methodological transparency.

The final data set consisted of 156 sources, including peer-reviewed publications, regulatory documents, policy reports, and professional guidance materials, all of which were subjected to thematic coding and interdisciplinary synthesis. To enhance methodological rigour, study selection and eligibility assessments were conducted collaboratively by all authors using the predefined inclusion and exclusion criteria. Any uncertainties regarding article eligibility, classification, or thematic relevance were discussed among the review team and resolved through consensus. This adjudication process helped ensure consistency and transparency throughout the screening, selection, and synthesis stages of the review.

2.8. Thematic Synthesis Methodology

Following study selection, the included sources underwent full-text review, thematic coding, and interdisciplinary synthesis. A thematic synthesis approach was employed, allowing identification of cross-domain patterns, emergent policy gaps, and unresolved ethical concerns. Each source was assessed for the presence of normative frameworks (e.g., fairness principles, transparency standards), empirical validation (e.g., model evaluation in diverse populations), and governance positioning (e.g., regulatory stance, oversight mechanisms). These sources were systematically classified into four thematic domains: ethical considerations, broader healthcare applications, implementation of diagnostic imaging, and governance, policy, and education.

In addition, where reported in the literature, indicators related to trust and real-world adoption, including clinician uptake, patient acceptance, calibration reliability, and concordance with expert judgement, were considered as part of the thematic synthesis. A supplementary data extraction matrix listing the included sources, evidence classifications, and thematic coding categories has been provided as Supplementary File S3 to support transparency, traceability, and reproducibility of the thematic synthesis.

2.9. Assessment of Evidence Types

Given the interdisciplinary nature of the included studies, a single quantitative risk-of-bias instrument was not applied. Instead, sources were categorised according to evidence type (empirical deployment study, clinical validation, technical method development, policy or regulatory guidance, and conceptual analysis).

Empirical deployment studies and regulatory documentation were prioritised when deriving governance implications, while conceptual and technical works informed ethical domain synthesis. This structured evidence categorisation strengthened interpretive consistency and supported alignment between proposed governance pillars and documented implementation challenges.

2.10. Methodological Considerations and Limitations

Several methodological limitations warrant consideration:

- a. **Language Scope:** The study included only sources in English, which may have excluded relevant regional insights.

- b. **Technological Rapid Evolution:** The review may not fully reflect the latest advances in foundation models and real-time AI deployment.
- c. **Stakeholder Diversity:** Perspectives from patients, frontline clinicians, and minority populations were limited in the peer-reviewed literature and were supplemented by grey literature and institutional reports where feasible.

The review was not prospectively registered because it was designed as a PRISMA-ScR-informed systematic survey focused on thematic synthesis, governance analysis, and interdisciplinary policy integration rather than quantitative intervention evaluation. Given the interdisciplinary and predominantly conceptual nature of the included literature, quantitative meta-analysis and formal effect-size synthesis were not appropriate.

Despite these limitations, the review methodology aligns with best practices in systematic and scoping reviews and reflects the interdisciplinary depth required to interrogate both the technical and normative dimensions of AI in medical imaging.

2.11. Example Search Strategy

To strengthen methodological transparency and reproducibility, an example Boolean search strategy used during the PRISMA-ScR-informed systematic survey process is provided below Box 1. Search queries were iteratively refined across databases to account for differences in indexing systems, controlled vocabularies, and platform-specific search functionalities. Complete database-specific search strategies, including database-adapted Boolean search strings, are provided in Supplementary Material S1 to support methodological transparency and reproducibility.

Additional keyword variations and database-specific adaptations were applied where appropriate to maximise interdisciplinary coverage across technical, ethical, regulatory, and clinical literature sources.

Box 1. Example Database: PubMed/MEDLINE.

("artificial intelligence" OR "machine learning" OR "deep learning" OR "large language models"
OR LLM OR CNN)
AND
("medical imaging" OR radiology OR diagnostics OR "diagnostic imaging")
AND
(ethics OR governance OR fairness OR explainability OR transparency OR accountability OR
privacy)
AND
(healthcare OR clinical OR hospital)

2.12. Synthesis Summary and Methodological Contribution

The structured methodology adopted in this survey enables a comprehensive and rigorous interrogation of ethical and governance challenges in AI-enabled diagnostic imaging. By integrating multi-source data, a clear classification scheme, and a robust synthesis strategy, this section forms the empirical foundation for the findings presented in Section 3.8. It directly supports the development of ethically grounded and policy-aware recommendations for the implementation of AI in healthcare. Importantly, this survey does not treat the reviewed literature as a collection of independent contributions, but as an interconnected body of work that reflects shared assumptions, recurring governance gaps, and divergent regulatory approaches. The synthesis presented in the following sections therefore emphasises cross-study comparison, thematic convergence, and points of tension between technical innovation and ethical oversight, rather than isolated summaries of individual publications.

3. Survey Results

The results presented in this section come from a comprehensive synthesis of scholarly and regulatory sources identified through a methodologically rigorous literature review protocol (see Section 2). The findings are mapped to the research questions described in Section 2 and provide a thematic distillation of current practices, ethical concerns, governance models, and stakeholder perspectives on AI integration in medical imaging.

For clarity, the results are presented in six linked parts: (1) search and selection outcomes, (2) characteristics of the included sources, (3) contemporary AI-enabled diagnostic imaging workflows, (4) ethical and governance risks identified across the literature, (5) stakeholder perspectives and trust-related findings, and (6) emerging governance models and synthesis of key findings. This structure distinguishes descriptive search outcomes from thematic findings and framework-relevant synthesis.

3.1. Search and Selection Summary

The literature search and screening process identified 642 records across the selected databases and supplementary sources. After removal of 98 duplicate records, 544 records underwent title and abstract screening. Following eligibility assessment, 213 full-text articles and documents were reviewed, resulting in the inclusion of 156 sources in the final synthesis. The included evidence comprised peer-reviewed publications, regulatory documents, policy reports, and professional guidance materials. The study selection process is summarised in Figure 4.

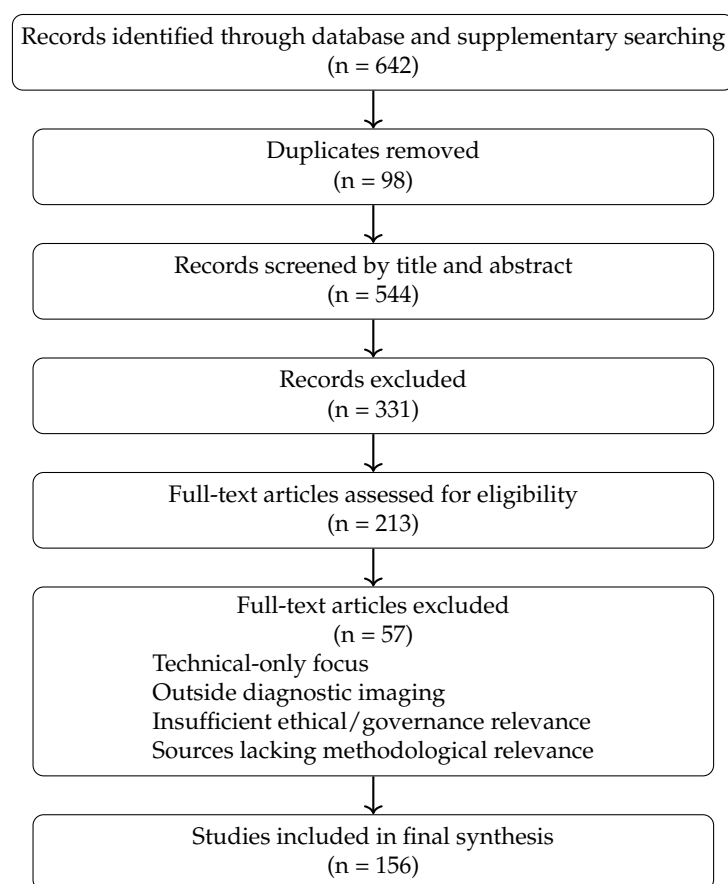


Figure 4. PRISMA-ScR flow diagram illustrating the study selection process for the systematic survey of sources published between 2018 and 2025.

3.2. Characteristics of Included Sources

The final dataset comprised 156 sources published between 2018 and 2025, reflecting the interdisciplinary nature of AI governance in diagnostic imaging. Sources included peer-reviewed journal articles, regulatory documents, policy reports, professional guidance materials, and institutional publications. For thematic synthesis, the included sources were categorised into four primary domains: ethical considerations (52 sources), broader healthcare applications (57 sources), diagnostic imaging implementation (43 sources), and governance, policy, and education (4 sources). Although some sources addressed multiple thematic areas, each was assigned to the domain representing its primary analytical focus. The distribution of included sources is presented in Table 1. To support transparent interpretation of the findings, evidence sources were analysed according to their primary contribution to the review. Empirical studies were used to identify reported implementation challenges, clinical outcomes, and deployment experiences. Regulatory documents, policy reports, and professional guidance materials informed the analysis of governance mechanisms, regulatory approaches, and accountability structures. Conceptual and normative publications primarily contributed to the synthesis of ethical principles, theoretical concerns, and emerging governance considerations. This evidence classification was maintained throughout the synthesis to ensure that policy recommendations and governance implications were interpreted in the context of the underlying evidence type.

3.3. Contemporary AI Workflows in Medical Imaging

3.3.1. AI-Integrated Diagnostic Pipelines

Analysis of the reviewed literature indicated that AI integration in diagnostic imaging most commonly occurs across four interconnected workflow stages: (i) *image acquisition*, (ii) *data preprocessing and harmonisation*, (iii) *AI model inference*, and (iv) *clinical decision support* [12,70]. Across the included studies, this modular workflow architecture emerged as the dominant model for incorporating AI into radiological practice. The literature further showed that convolutional neural networks (CNNs), transformers, and more recently large language models (LLMs) were the technologies most frequently applied within these workflows. These systems were commonly used for tasks including anomaly detection, segmentation, disease classification, risk stratification, and automated reporting.

3.3.2. Data Harmonisation and Preprocessing Techniques

The reviewed literature consistently identified data harmonisation and preprocessing as critical prerequisites for reliable AI performance across clinical settings. Commonly reported approaches included contrast normalisation, artefact removal, and anatomical labelling [44,70,109–117]. These steps reduce scanner variability and improve generalisability across institutions, key to multicentre AI deployment.

3.3.3. AI Inference and Clinical Integration

Across the included studies, AI inference systems were most commonly applied to disease classification, risk stratification, and treatment prioritisation tasks. Several studies also reported increasing use of multimodal approaches [13,101]. LLMs are increasingly used to synthesise multimodal input (e.g., images + EHR notes), generating clinically actionable insights. However, this introduces complexities in interpretability and legal accountability [34,71].

3.3.4. Systemic Feedback and Lifecycle Integration

The reviewed studies indicated a growing trend towards integrating AI systems with radiology infrastructures such as PACS and RIS. Feedback mechanisms from clinicians

support continuous learning, but pose challenges regarding data provenance, system auditability, and over-reliance [10,14,112,142,144].

3.4. *Ethical and Governance Risks Across Workflows*

3.4.1. Opacity and Interpretability Deficits

The reviewed literature consistently identified the black-box nature of deep learning systems as a prominent challenge, undermining clinician trust and the post hoc justification of AI decisions [70,73]. Studies further highlighted implications for patient consent, shared decision-making, and medico-legal responsibility [40,144].

3.4.2. Bias and Underrepresentation

The reviewed studies consistently reported algorithmic bias arising from unbalanced or non-representative datasets [69,99]. Disparate performance in subgroups, particularly among under-represented populations and rare pathologies, risks perpetuating inequities in healthcare provision [11,26].

3.4.3. Governance Fragmentation and Lifecycle Blind Spots

The reviewed literature identified a lack of consensus regarding risk classification, post-market surveillance, and model update mechanisms, resulting in persistent regulatory gaps [101,103]. Studies further indicated that few jurisdictions adequately address the dynamic nature of AI systems, particularly those capable of real-time learning, continual adaptation, or performance drift [70,107].

3.4.4. Data Security and Privacy Exposure

The reviewed literature highlighted data security and privacy risks associated with the large volumes of sensitive health data required by AI systems, including concerns relating to reidentification, consent, and regulatory compliance [27,44]. Studies further indicated that although federated learning and encryption-based approaches offer potential mitigation pathways, their adoption in operational clinical environments remains limited [70].

3.5. *Stakeholder Perspectives on Ethical AI*

Stakeholders express various perceptions based on professional roles.

- **Clinicians and Radiologists:** Supportive of AI's diagnostic potential but cautious due to trust and accountability concerns [13,149].
- **Patients:** Recognise benefits of speed and personalisation but express concern about the misuse of data and the opacity of decisions [11,41].
- **Developers:** Focus on technical efficacy and data access, often underestimating ethical downstream effects [46,74].

This misalignment requires the participation of multiple stakeholders in AI design and deployment [73,97].

3.6. *Trust as a Measurable Dimension of Clinical AI Governance*

Although trust is frequently invoked in discussions of artificial intelligence (AI) governance in healthcare, it is often treated as a broad normative aspiration rather than as a measurable implementation outcome. In AI-enabled medical imaging, this creates an important translational gap. A system may satisfy high-level expectations relating to fairness, transparency, privacy, and accountability, yet still fail to achieve meaningful clinical adoption if clinicians do not trust its outputs, patients do not accept its use, or institutions cannot verify its reliability under real-world conditions [41,109,117]. Trust should therefore

be understood not only as an ethical objective, but also as an operational dimension of governance that can be assessed through empirical indicators across the AI lifecycle.

Within diagnostic imaging, trust is shaped by both technical and sociotechnical conditions. At the technical level, clinicians are more likely to rely on AI systems when predicted probabilities are well calibrated, when performance remains stable across settings, and when outputs are consistent with expert judgement in clinically meaningful tasks. At the sociotechnical level, trust is influenced by interpretability, accountability, and stakeholder understanding of AI-supported decisions [40,74,116]. These factors have been widely identified as barriers to trustworthy deployment, particularly where systems remain opaque or poorly integrated into clinical workflows.

To address this gap, trust can be integrated into lifecycle evaluation through a set of measurable indicators that connect governance principles to observable outcomes. During model development and validation, calibration measures can assess whether predicted probabilities align with true outcome frequencies, thereby indicating whether the system communicates confidence in a clinically reliable manner. During validation and deployment, radiologist–AI concordance can provide an empirical signal of alignment between model outputs and expert interpretation. During clinical use and post-deployment monitoring, patient acceptance rates, clinician uptake patterns, and override behaviour can help determine whether the system is trusted in practice rather than merely approved in principle [89,109].

These indicators should not be interpreted in isolation. Calibration without fairness may reinforce inequitable trust, while concordance without interpretability may conceal automation bias. For this reason, trust should be conceptualised as a cross-cutting governance outcome emerging from the interaction of ethical and regulatory dimensions identified in this review, including fairness, transparency, accountability, and privacy [48,69]. Trust therefore functions as an evaluative layer that determines whether governance mechanisms are effective in real-world clinical contexts.

From a governance perspective, integrating trust metrics strengthens the practical relevance of policy recommendations. Rather than recommending transparency, fairness auditing, or post-market monitoring only as normative safeguards, institutions can attach measurable trust-related indicators to each stage of implementation. This enables a shift from principle-based governance toward evidence-informed governance capable of evaluating whether AI systems are not only technically functional and legally compliant, but also clinically credible and socially legitimate.

In this sense, trust quantification does not replace ethical governance; rather, it provides the missing bridge between governance design and real-world adoption. For AI in medical imaging to be deployed responsibly, governance frameworks must demonstrate not only what safeguards exist, but also whether those safeguards produce trustworthiness that is visible, measurable, and sustainable in practice.

3.7. Emerging Governance Models

Several jurisdictions have introduced AI-specific regulatory or guidance frameworks. The European Union's AI Act classifies medical AI as high-risk and mandates conformity assessments and post-market monitoring. The U.S. FDA regulates AI-enabled medical devices through Software as a Medical Device (SaMD) pathways, increasingly incorporating adaptive oversight mechanisms. The United Kingdom and Canada have emphasised lifecycle governance and algorithmic transparency, while China has introduced targeted algorithm regulations for clinical and consumer-facing systems.

Despite structural differences, these models share common themes: risk-based classification, documentation requirements, and increasing attention to post-deployment monitor-

ing. However, operationalisation of adaptive oversight for continuously learning systems remains underdeveloped across jurisdictions. The European Union (EU), United States (US), and Australia were selected for detailed analysis because they represent three influential and contrasting approaches to AI governance in healthcare. The EU provides a comprehensive risk-based regulatory framework through the AI Act, the US adopts a sector-specific and lifecycle-oriented regulatory model centred on FDA oversight, and Australia represents an emerging governance environment that combines existing regulatory mechanisms with voluntary ethical principles. Together, these jurisdictions offer a useful comparison of mature, evolving, and hybrid governance approaches, while providing insights relevant to the responsible implementation of AI in healthcare systems internationally.

3.7.1. The European Union (EU)

The EU leads globally with its proposed *AI Act 2024*, a landmark regulatory initiative that introduces a risk-based classification framework for artificial intelligence applications. In this approach, AI systems are classified into unacceptable, high, limited, and minimal risk levels, with healthcare applications—especially those that involve diagnosis or clinical decision support—typically falling into the high-risk category [13,107]. This classification triggers stringent requirements, including mandatory conformity assessments, algorithmic transparency obligations, data quality standards, cybersecurity safeguards, and provisions for human oversight and intervention [1–10,15,26,59,60,64,69,126–132].

A defining feature of the EU AI Act 2024 is its emphasis on upstream risk mitigation. High-risk systems must comply with technical documentation rules that enable traceability and reproducibility, including logging mechanisms and lifecycle performance reporting. The Act also mandates post-market surveillance obligations and empowers national supervisory authorities with enforcement capabilities in EU member states [62,65]. This comprehensive framework represents a shift away from reactive and harm-based regulation to anticipatory governance based on system design, deployment context, and intended functionality [61,63]. However, implementation challenges persist. The different levels of digital maturity, regulatory capacity, and healthcare infrastructure in EU member states threaten to create inconsistencies in enforcement and interpretation [67,68]. Furthermore, the AI Act currently lacks explicit mechanisms for dealing with adaptive or self-improving algorithms, such as continual learning systems or foundation models, that evolve after market approval. This regulatory gap raises concerns about auditability, accountability, and model drift in real-world clinical settings [72,107,108].

In addition, the Act does not yet provide detailed technical criteria for explainability, fairness, or robustness, leaving open questions about how these principles should be operationalised in clinical workflows [34,73]. Stakeholders have also expressed concerns about the compliance costs of smaller AI developers and healthcare institutions, who may lack the legal and technical resources required to meet these obligations [70,100]. Despite these limitations, the EU AI Act 2024 sets an important precedent in establishing a unified, legally binding framework for ethical AI governance. Its influence is already observable in other jurisdictions, prompting parallel efforts in Canada, Brazil, and parts of Southeast Asia to design risk-based AI regulation [11,13,19,44,70]. If effectively implemented and continuously updated to account for technological advancements, the Act could serve as a global benchmark for responsible AI deployment in healthcare and beyond.

3.7.2. The United States (US)

The US adopts a lifecycle-centered approach to AI governance in healthcare, regulated primarily through the Food and Drug Administration's (FDA) *Software as a Medical Device (SaMD)* framework. This model emphasises real-world performance monitoring, benefit-

risk assessments, and update traceability throughout the product lifecycle. Rather than approving AI systems based solely on static pre-market submissions, the FDA encourages adaptive oversight through its *AI/ML-Based SaMD Action Plan*, which lays out principles for algorithm transparency, performance evaluation, and a total product lifecycle (TPLC) approach [16,17,103,104]. Central to this framework is the requirement for manufacturers to demonstrate analytical and clinical validation through real-world evidence (RWE), including post-market surveillance data, to support safety and effectiveness in various clinical environments [19,20]. The model also promotes transparency around algorithmic changes, particularly for systems employing machine learning (ML) or deep learning (DL) techniques that may evolve after deployment. The concept of a “Predetermined Change Control Plan” (PCCP) has been introduced to allow for regulated, pre-approved updates without the need for full resubmission, an essential mechanism for supporting continuous learning systems [24,25].

Despite these innovations, several regulatory gaps remain. The FDA currently lacks a dedicated AI-specific regulatory category different from traditional SaMDs, which can lead to interpretive inconsistencies, especially in cases involving generative models or complex multimodal inference engines [18,22]. This ambiguity complicates clinical integration, particularly when AI systems serve as autonomous agents in classification, diagnosis, or therapeutic decision-making [21,23]. Furthermore, concerns persist around the scalability and timeliness of FDA approvals, with critiques citing long review timelines and high evidentiary burdens that disproportionately affect smaller developers and startups [17,23]. The voluntary nature of some transparency recommendations, such as sharing model performance between demographic subgroups or disclosing the explainability limitations, also raises concerns about enforceability and accountability [16,24].

Unlike the centralised approach of the European Union, the US relies on a more fragmented and market-driven regulatory environment, often mediated by state laws, private sector ethics boards, and institutional review bodies. This fragmentation leads to variable enforcement between jurisdictions and institutions, undermining consistent ethical safeguards [19,25]. Despite these challenges, the FDA’s iterative evidence-based model holds promise for aligning innovation with safety and efficacy, particularly if supported by clearer federal legislation that examines the unique risks posed by adaptive AI systems. Ongoing legislative discussions, including proposals for an AI Bill of Rights and updates to HIPAA to account for AI-driven analytics, suggest a growing political will to strengthen regulatory clarity and public trust [17,21,23,25,103].

3.7.3. Australia

Australia’s approach to AI governance in healthcare is characterised by a progressive, innovation-friendly orientation. The Australian Health Practitioner and Regulation Agency (AHPRA), Australia’s principal medical regulator, has embraced regulatory agility through the use of *regulatory policies* and pilot programmes that allow AI developers to test technologies in a controlled environment without facing full regulatory burdens. This has enabled experimentation with high-risk adaptive AI systems while mitigating immediate legal risk [3,6,70].

Complementing this technical flexibility, Australia has developed a set of voluntary ethical principles in AI under the auspices of the Department of Industry, Science and Resources. These principles emphasise values such as human-centered design, explainability, privacy, and accountability, in order to guide ethical AI development across sectors, including healthcare [4,46]. Furthermore, the Australian Alliance for Artificial Intelligence in Healthcare (AAAIH) has issued sector-specific frameworks to support AI adoption in clinical settings while maintaining transparency and public trust [5,10]. Despite these

forward-leaning initiatives, significant regulatory gaps remain. One of the most pressing challenges is the lack of binding AI-specific legislation governing medical AI tools. Current approval pathways are still mainly grounded in traditional medical device frameworks, which are not well suited to address the unique risks associated with complex multimodal platforms [1,2,7]. This has led to inconsistency in the oversight practices between institutions and states, and some rely on institutional ethics boards or vendor self-assessment to evaluate AI systems, often without formalised national guidance [8,9]. This raises serious concerns about the long-term accountability, safety, and fairness of AI once integrated into routine clinical workflows. Current frameworks also do not assess demographic fairness, cultural safety, and accessibility for Aboriginal and Torres Strait Islander populations, exacerbating existing healthcare inequalities [26,70]. However, Australia is uniquely positioned to become a regional leader in AI governance. Its robust public health infrastructure and early engagement with AI ethics provide a foundation for crafting agile and context-sensitive regulatory frameworks [31,48,49,107].

Efforts such as the Australian Government's Digital Health Strategy and the AI Action Plan signal the growing political recognition of these challenges. However, translating voluntary principles into enforceable sector-specific regulations will be critical to avoiding a governance vacuum as AI becomes ubiquitous in diagnostic imaging, triage, and clinical decision support systems [13,44,70,74]. In federated systems such as Australia, regulatory responsibilities are distributed across national and state authorities, creating additional coordination challenges [14,35,50,70].

Other jurisdictions, including the United Kingdom, Canada, and China, have introduced guidance frameworks emphasising transparency, lifecycle oversight, and algorithm accountability, though operational standards for adaptive systems remain in development [31,32,35,46,48,50].

3.8. The Need for a Structured AI Governance Model

Rapid integration of artificial intelligence (AI) into medical imaging has catalysed significant advances in diagnostic precision, workflow optimisation, and clinical decision support. However, these innovations also introduce profound ethical dilemmas and regulatory ambiguities that existing governance structures are ill-equipped to handle. In the absence of comprehensive oversight, the deployment of AI risks undermining public trust, exacerbating healthcare disparities, and enabling opaque decision-making that challenges traditional notions of clinical accountability.

As AI systems transition from static tools to adaptive, continuously learning entities, the regulatory demands placed on them increase exponentially. These technologies often operate in high-stakes environments, such as radiological triage, cancer screening, and surgical planning, where the cost of error can be life-changing. Moreover, the "black box" nature of many deep learning architectures complicates transparency and explainability, two essential prerequisites for ethical and safe implementation. Consequently, governance must evolve from fragmented, reactive measures into a proactive, lifecycle-aware framework that can anticipate and mitigate both known and emerging risks.

A structured and context-sensitive governance model for AI in Australian medical imaging should reflect not only international best practices but also the unique sociotechnical, cultural and legal dimensions of the local healthcare system. Drawing on comparative insights from global frameworks and recent regulatory experiments, such a model must rest on the following foundational pillars [6,13,70,101,103,104]:

- **Regulatory Standardisation:** AI governance must be codified through enforceable sector-specific legislation. These laws should define approval pathways for medical AI tools, establish pre-market testing protocols, mandate real-world performance

monitoring, and incorporate mechanisms for dynamic reevaluation, especially for models that evolve over time. Regulatory alignment with international norms (for example, the EU AI Act 2024 or the FDA's SaMD framework) can facilitate global interoperability, but must be adapted to local contexts and ethical expectations.

- **Bias Mitigation and Equity Assurance:** Structural bias in training data, model architecture, and clinical deployment perpetuates health inequities, disproportionately affecting historically marginalised groups. Governance should require fairness audits, demographic performance stratification, and equity impact assessments throughout the AI lifecycle. In addition, funding mechanisms and incentives for data sharing must be created to support the inclusion of under-represented populations in clinical data sets.
- **Explainability and Clinical Transparency:** Regulatory bodies must enforce explainability as a legal requirement for high-risk AI systems. This includes mandates for interpretable output, model rationale reporting, and user-level documentation that clinicians and patients can understand. The explanation should extend beyond technical traceability to include meaningful and context-relevant communication that supports informed consent and shared decision making.
- **Accountability and Legal Clarity:** Legal responsibility for AI-induced harm must be clearly allocated throughout the AI value chain. This includes defining the liability between developers, healthcare providers, institutions, and vendors. Governance frameworks should require documentation of AI decision-making processes (e.g., audit trails), robust incident reporting mechanisms, and third-party validation of safety claims. Without these measures, accountability becomes diffused, hindering redress, and eroding public trust.
- **Adaptive and Participatory Oversight:** Given the pace of technological innovation, static regulatory models will quickly become obsolete. A future-proof governance model must incorporate mechanisms for continuous learning, stakeholder consultation, and adaptive policy revision. Public and professional involvement, through citizen juries, ethics boards, and clinical feedback loops, should be embedded into the oversight infrastructure to ensure that evolving standards reflect societal values and frontline realities.

This governance vision must be co-designed through interdisciplinary collaboration, involving not only clinicians and technologists, but also patients, ethicists, policymakers, and legal experts. Such inclusive design processes are crucial for ensuring that ethical safeguards are embedded from the outset rather than retrofitted after deployment. In addition, Australia is well positioned to contribute to leadership in this area by developing a hybrid governance model that bridges ethical principles with enforceable regulation, tailored to its healthcare infrastructure, policy landscape, and digital maturity.

Table 7 presents a comparative overview of emerging AI governance models across major jurisdictions and international organisations. The table highlights differences in regulatory approaches, enforceability, governance mechanisms, and implementation challenges, providing a structured basis for understanding how various healthcare systems are responding to ethical, legal, and operational challenges associated with AI-enabled medical technologies [46,70,103,104,107]. This comparison informs the development of the proposed governance framework by identifying common strengths, recurring limitations, and opportunities for regulatory harmonisation across healthcare AI governance systems.

Table 7 demonstrates that while regulatory approaches differ substantially across jurisdictions, common challenges persist regarding adaptive AI oversight, accountability mechanisms, transparency, and post-deployment governance.

Table 7. Comparative Overview of Emerging AI Governance Models in Healthcare Across Major Jurisdictions and International Organisations.

Jurisdiction/Body	Regulatory Model	Enforceability	Core Governance Features	Key Limitations and Gaps
European Union (EU)	AI Act (risk-based)	Legally binding	Conformity checks, transparency, data traceability, post-market surveillance	No standards for adaptive AI; country-level variation; high compliance burden
United States (FDA)	Lifecycle-centered (SaMD)	Medium (sectoral)	Total Product Lifecycle (TPLC), real-world evidence, model change protocols (PCCP)	No dedicated AI category; slow approvals; fragmented oversight
Australia (TGA)	Innovation + voluntary ethics	Low–Medium (hybrid)	Policies, AI Ethics Principles, pilot testing pathways	No binding AI law; weak post-market monitoring; fairness not enforced
United Kingdom (UK)	Sector-led soft governance	Low (principles-based)	MHRA oversight, explainability, public trust initiatives	No central AI body; unclear accountability; legal ambiguity
Canada (Health Canada)	Draft AI Act (AIDA)	Low (pending)	Bias audits, non-discrimination, clinical validation	Early stage; uncertain scalability; overlaps in jurisdiction
Singapore (PDPC)	Co-regulatory policies	Low (voluntary)	Explainability, robustness, flexible policies	No enforcement power; redress unclear for high-risk use
China (NHC/Cyber Law)	State-centered mandates	High (state-controlled)	Algorithm registration, performance audits, data localisation	Limited transparency; opaque approvals; weak civil oversight
Japan (PMDA)	Post-market focused	Medium	Algorithm traceability, safety-first regulation	Slow on LLMs and generative AI; limited stakeholder feedback
WHO/OECD	Normative guidance	None	Rights-based ethics, transparency, sustainability	Advisory only; no enforcement; limited uptake in LMICs

3.9. Synthesis of Key Findings

Thematic synthesis of the 156 included sources revealed several recurring patterns across the AI lifecycle. Although technological capabilities continue to advance rapidly, the literature consistently identified ethical, governance, and implementation challenges that may limit safe, equitable, and trustworthy deployment in clinical practice. The principal findings are summarised below.

- AI workflows in medical imaging are increasingly modular and multimodal, but remain constrained by ongoing challenges relating to explainability, transparency, and accountability.
- Ethical risks, including algorithmic bias, opacity, and privacy exposure, are widely documented but remain insufficiently addressed in many real-world clinical deployments.
- Stakeholders express divergent priorities and expectations, underscoring the importance of participatory, interdisciplinary, and lifecycle-aware governance approaches.
- Regulatory strategies remain fragmented, predominantly premarket-oriented, and insufficiently equipped to support continuous oversight of adaptive and continuously learning AI systems.
- Trust emerged as a cross-cutting governance concern, with the literature highlighting the need for measurable indicators such as calibration reliability, clinician–AI concordance, patient acceptance, and real-world adoption to support responsible implementation.

These findings reinforce the need for robust, enforceable, and interdisciplinary governance frameworks capable of balancing innovation with trust, safety, accountability, and patient-centred care.

To support analytical synthesis across the reviewed literature, ethical and governance challenges were organised according to the AI system lifecycle. This lifecycle-centric taxonomy enables systematic comparison of risks, governance responses, and recurring limitations reported across studies, providing a structured foundation for the thematic analysis and governance recommendations presented in the following sections.

Rather than treating ethical challenges as isolated concerns, the taxonomy highlights how governance gaps recur and evolve across different stages of the AI lifecycle. This structure is used throughout the remainder of the survey to guide comparative analysis and to inform the development of governance recommendations.

4. Integrated Findings and Synthesis

The systematic analysis of 156 scholarly and regulatory sources highlights deeply embedded and intersecting ethical and governance challenges surrounding the adoption of Artificial Intelligence (AI) in medical imaging and diagnostics. Although AI-driven technologies demonstrate growing promise in improving diagnostic accuracy, workflow efficiency, and personalised care, significant unresolved vulnerabilities persist in multiple dimensions. These include enduring deficits in algorithmic transparency and interpretability, structural biases arising from unrepresentative training datasets that risk exacerbating healthcare disparities, insufficient privacy-preserving mechanisms for handling sensitive medical data, and fragmented regulatory frameworks ill-equipped to govern rapidly evolving AI systems. The increasing deployment of sophisticated AI architecture, such as foundation models, generative systems, large multimodal models, and continuously learning algorithms, further intensifies these challenges by introducing adaptive behaviours, real-time model drift, and accountability ambiguities that transcend the capabilities of existing governance structures. Despite emerging governance efforts in select jurisdictions, current regulatory models often remain reactive, pre-market focused, and lack enforceable lifecycle oversight capable of managing the complex sociotechnical risks posed by AI-enabled diagnostic systems in diverse clinical environments [70,72,101,122]. Algorithmic bias remains a central concern, as many AI models are trained in data sets that do not adequately represent diverse populations, leading to unequal diagnostic outcomes between demographic groups [23,69,98]. The opacity of deep learning systems, often described as “black-box” models, continues to undermine trust and informed clinical decision making, despite advances in explainability tools such as saliency maps and local surrogate models [25,73,126,127].

A cross-cutting finding of this review is that while trust is frequently invoked as a central objective of ethical and regulatory frameworks for clinical AI, it is rarely operationalised through measurable or auditable indicators. Across the reviewed literature, trust is often assumed to emerge from model performance, transparency mechanisms, or regulatory compliance, yet few studies provide concrete approaches for evaluating trust in practice. Where measurable proxies are discussed, they tend to include factors such as model calibration reliability, clinician–AI concordance, patterns of clinical adoption or override behaviour, and patient acceptance in real-world contexts. This gap highlights a critical limitation in current governance approaches, where trust remains a normative aspiration rather than an empirically grounded dimension of AI system evaluation.

Legal accountability remains poorly defined in scenarios where autonomous AI systems contribute to clinical errors, raising unresolved questions about the allocation of responsibility among developers, vendors, and healthcare providers [100,103,145]. In parallel, data privacy concerns persist, driven by the scale and sensitivity of medical data aggregated for AI training. Existing consent models and anonymisation frameworks often lag behind technical developments, which requires more robust privacy preservation

approaches, such as federated learning and secure multiparty computation [27,44]. At the governance level, regulatory fragmentation remains a global obstacle, with divergent standards and inconsistent oversight mechanisms between jurisdictions such as the FDA (USA), EMA (Europe), and TGA (Australia) [45,104,146]. Although some progress has been made toward developing international regulatory frameworks, these efforts often struggle to keep up with the rapidly evolving AI capabilities [46,70,77]. Crucially, these findings highlight a growing temporal gap between emerging ethical challenges and the slower institutional development of governance capacity. This is particularly relevant in the Australian healthcare context, where AI-driven imaging technologies are advancing, but governance mechanisms are still in the early stages [79,101]. Addressing these gaps requires not only ongoing interdisciplinary dialogue, but the formulation of adaptive, context-sensitive policy frameworks that prioritise fairness auditing, transparent accountability, and participatory oversight structures [31,86,108].

To facilitate interpretation of the integrated findings, Figure 5 presents a conceptual synthesis of the principal ethical and governance challenges identified across the reviewed literature. The figure illustrates how recurring concerns relating to bias and equity, explainability and transparency, privacy and consent, accountability and liability, and regulatory fragmentation collectively influence stakeholder trust and the responsible implementation of AI in medical imaging and diagnostics.

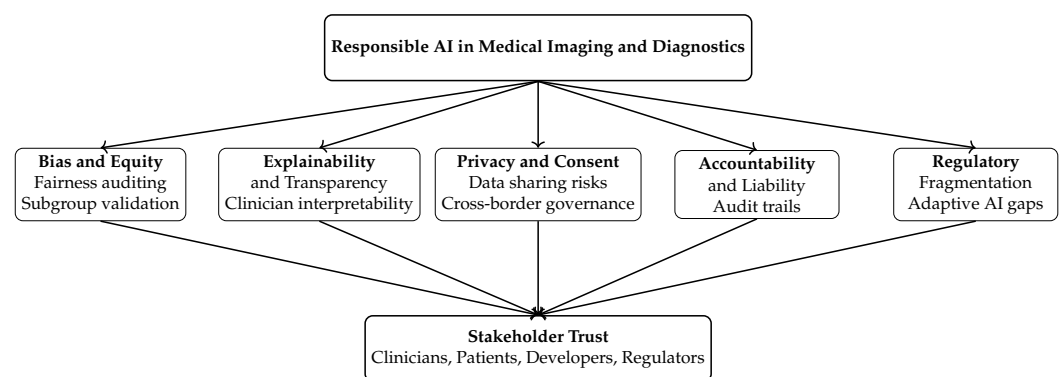


Figure 5. Conceptual synthesis of the principal ethical and governance challenges identified across the reviewed literature and their collective influence on stakeholder trust and the responsible implementation of AI in medical imaging and diagnostics.

Figure 5 illustrates that stakeholder trust emerges from the interaction of multiple ethical and governance domains rather than any single challenge in isolation, reinforcing the need for integrated lifecycle governance approaches.

4.1. AI Ethics vs. Technical Safety: Conceptual Divergence

In the reviewed literature, the terms *ethics*, *governance*, and *safety* are often conflated, although they serve distinct functions. Technical safety relates to robustness and performance; ethics considers fairness, autonomy, and rights; and governance refers to policy structures that ensure ethical compliance [13,70,101]. This conceptual ambiguity often results in fragmented strategies that neglect the full scope of responsible AI integration. For example, explainability methods are frequently promoted for clinical acceptance, but insufficient attention is paid to their ethical implications in decision-making [24,34,73].

4.2. Data Privacy and Consent in AI Imaging Systems

AI models in medical imaging are predominantly trained on large datasets containing sensitive patient information. However, a significant portion of the literature omits or marginalises privacy-preserving mechanisms such as federated learning, differential

privacy, or secure multiparty computation [27,44]. Moreover, while HIPAA and GDPR are cited as guiding standards, implementation in jurisdictions remains inconsistent [32,70]. The use of cloud-based systems further exacerbates these concerns, especially when data is stored and processed across international borders with varying legal protections [13,101].

4.3. Algorithmic Bias: The Unresolved Equity Challenge

Numerous articles reported bias in the output of the AI model due to the under-representation of minority populations in training data [69,98,99]. This includes disparities in image quality, scanner types, and annotation accuracy in geographies. Despite growing awareness, systematic fairness auditing and demographic performance benchmarking remain largely absent from commercial and academic systems. Some studies advocated for debiasing techniques, but these approaches lack consensus on metrics and standards [26,34,97].

4.4. Opacity and the Need for Explainable AI (XAI)

Deep learning models used in imaging diagnostics often function as black-box systems. This opacity undermines trust and obstructs the legal and clinical validation of AI decisions [71,72,127]. XAI techniques such as saliency maps, Shapley values, and counterfactual explanations have been proposed, yet they are inconsistently adopted and insufficiently evaluated in medical contexts [24,31,73]. The literature reveals a tension between model performance and interpretability, with most clinical stakeholders favouring transparency over marginal gains in accuracy.

4.5. Governance Models and Global Fragmentation

The governance of AI in medical imaging lacks uniformity between jurisdictions. Our analysis revealed divergent policy frameworks, ranging from voluntary ethical principles (e.g., Australia's AI Ethics Principles) to formal regulatory instruments (e.g., the EU AI Act 2024, FDA SaMD) [70,104,107]. Table 7 summarises these frameworks. In particular, Australia currently lacks enforceable legislation specifically addressing AI in healthcare protocols, making legal harmonisation, standards development, and multi-stakeholder oversight an urgent need [46,101].

4.6. Clinical and Stakeholder Perceptions

Clinicians value the ability of AI to accelerate diagnosis and improve decision making, but express concerns about its reliability and the risk of deskilling [41,42,139]. Patients want faster and more accurate diagnoses, but remain sceptical of data misuse and lack of transparency [44,58]. Developers prioritise model accuracy and data access, but often overlook ethical design and regulatory fit. These perspectives underscore the misalignment between technical advancement and societal readiness [11,13,46,70,79,98,100,107–117,135–138].

4.7. Discussion

Our systematic survey identifies six critical and interlinked findings that reflect both thematic convergence and persistent fragmentation across ethical, technical, and governance domains in AI-driven medical imaging. These insights underscore the urgent need for a comprehensive and enforceable governance framework geared to national contexts like Australia, while informed by global best practices [71–73,76–96,118–125,133,134,140,141,150–153].

- **Conceptual Misalignment Between Ethics, Safety, and Governance.** Across jurisdictions, ethical concerns (for example, fairness, consent), technical safety (for example, robustness), and governance (for example, regulation, liability) are frequently conflated, resulting in misaligned priorities, patchwork accountability, and unclear

deployment standards [13,69,70,101]. Lack of definitional clarity inhibits stakeholder collaboration and weakens policy coherence.

- **Privacy Gaps and Cross-Border Data Vulnerabilities.** Despite increasing use of cloud-based infrastructure and transnational training pipelines, privacy-preserving techniques (e.g., federated learning, homomorphic encryption) remain underused or poorly regulated [27,44,120]. International divergence in data protection laws further complicates cross-jurisdictional compliance and undermines trust in AI systems [11,13,32,70,98,100,107,108,135–138].
- **Systemic Algorithmic Bias and the Absence of Equity Auditing.** Disparities in model performance across demographic subgroups, often resulting from non-representative training data, persist as one of the most widely cited and least addressed concerns in the literature [26,69,99]. Most regulatory frameworks do not mandate fairness benchmarking, leaving bias mitigation dependent on voluntary developer practices [31,100,154].
- **Explainability Remains a Bottleneck Despite Technical Progress.** While saliency maps, SHAP/LIME, and counterfactual explanations offer interpretability pathways, their clinical utility and reliability remain contested. Moreover, explanation tools have not yet been standardised or required by most governance regimes, posing risks to informed consent, clinician trust, and medicolegal defence [34,72,73].
- **Regulatory Fragmentation and Lifecycle Blind Spots.** Most jurisdictions lack comprehensive and enforceable governance structures that cover the entire AI lifecycle, including real-time learning, adaptive updates, and post-deployment monitoring [103,107]. Table 7 highlights the spectrum of maturity in governance, with only the EU proposing legally binding lifecycle oversight through its AI Act [104].
- **Stakeholder Priorities are Misaligned and Under-represented.** Clinicians emphasise interpretability and workflow integration; patients prioritise consent and data control; and developers focus on precision and scalability, yet few governance models institutionalise mechanisms for participatory co-design [41,46,58,155]. This misalignment exacerbates socioethical disconnects and delays real-world adoption.

Collectively, these findings reaffirm the necessity of a structured and interoperable ethical governance model outlined in Section 1.4 that incorporates enforceable standards, equity auditing, explainability benchmarks, and multistakeholder oversight. This model is essential to ensure that AI innovations in Australian healthcare are not only technically robust but also ethically sound, legally accountable, and socially trustworthy [13,70,100,156].

An important implication of this review is that trust in clinical AI is predominantly treated as a normative or aspirational concept rather than a measurable property of system performance and governance. While ethical and regulatory frameworks consistently emphasise the importance of trust, they rarely provide mechanisms for assessing it across different stages of the AI lifecycle. This study highlights the need to conceptualise trust as an evaluative dimension that can be examined through indicators such as calibration reliability, clinician–AI concordance, patterns of adoption and override behaviour, and patient acceptance in real-world clinical settings. Embedding such measurable trust indicators within governance frameworks can strengthen policy relevance, support implementation readiness, and enable more accountable and transparent integration of AI in diagnostic practice.

4.8. Implications for Digital Health Systems

The findings underscore that AI-enabled diagnostic imaging should be governed as part of an integrated digital health ecosystem rather than as isolated medical devices.

Governance mechanisms must therefore align with electronic health record interoperability standards, institutional risk management structures, and national digital health strategies.

Lifecycle monitoring, subgroup performance transparency, and adaptive regulatory pathways are not solely technical safeguards; they are institutional capabilities that determine whether digital health systems can sustain public trust and clinical reliability. Without coordinated system-level oversight, technological innovation risks outpacing governance maturity across interconnected digital infrastructures.

4.9. Transferability Beyond the Australian Context

Although the proposed conceptual governance framework is illustrated using the Australian regulatory environment, its structural components are intentionally jurisdiction-agnostic. The lifecycle-centric approach, bias-auditing mechanisms, subgroup-performance reporting, human-in-the-loop oversight, and post-deployment monitoring requirements can be adapted across federated and centralised health systems.

Jurisdiction-specific elements such as privacy statutes or medical device pathways may vary, but the underlying governance architecture remains applicable to digital health ecosystems facing similar challenges of regulatory fragmentation, adaptive AI systems, and cross-institutional data flows.

Although the framework is informed by Australian healthcare governance arrangements and regulatory mechanisms, its underlying governance principles are derived from a broader international evidence base, including guidance from the World Health Organization, OECD AI Principles, FDA Good Machine Learning Practice guidance, IMDRF recommendations, and the European Union AI Act. Consequently, the framework is intended as a transferable conceptual structure that can be adapted to different healthcare systems while incorporating jurisdiction-specific legal, regulatory, organisational, and cultural requirements.

5. Conclusions and Future Research Directions

The rapid integration of AI into diagnostic imaging presents not only technical innovation but also increasing governance complexity across digital health ecosystems. This review synthesised 156 studies to identify recurring ethical risks and regulatory gaps across the AI lifecycle. The reviewed literature consistently highlighted challenges relating to explainability, algorithmic bias, accountability, privacy protection, stakeholder inclusion, and limitations in existing regulatory approaches for adaptive AI systems.

These findings directly reflect the thematic patterns identified throughout the Results section, where recurring governance gaps, stakeholder concerns, lifecycle management limitations, trust-related challenges, and regulatory fragmentation emerged consistently across the reviewed evidence base. Building directly on the findings synthesised from the reviewed literature, our lifecycle-centric governance framework translates widely recognised ethical principles into practical oversight mechanisms spanning development, deployment, monitoring, and adaptive updating. Although illustrated through the Australian context, the framework is transferable to federated and multi-jurisdictional health systems facing similar governance challenges.

The governance framework and policy recommendations presented in this review represent the authors' synthesis and interpretation of evidence derived from empirical studies, regulatory documents, policy reports, and conceptual analyses rather than direct findings reported uniformly across all included sources. As a literature-derived conceptual framework, the proposed governance framework provides a structured foundation for future implementation and evaluation. Future research should assess its applicability, usability, and effectiveness across diverse clinical, organisational, and regulatory settings.

Future digital health governance should prioritise post-deployment surveillance, sub-group performance transparency, adaptive regulatory mechanisms for continuously learning systems, and measurable trust indicators, including calibration reliability, clinician–AI concordance, adoption and override behaviour, patient acceptance, and post-deployment monitoring signals. Without such institutional readiness, technological advancement may outpace the safeguards required for equitable and trustworthy clinical AI integration.

5.1. Persistent Limitations Hindering Ethical AI Integration

Despite growing scholarly attention and emerging regulatory initiatives, several structural and interrelated limitations continue to hinder the ethical integration of AI in medical imaging.

- **Absence of Clinically Grounded Explainability Standards:** While explainable AI (XAI) has advanced methodologically, most current techniques remain insufficient for clinical decision-making settings. Tools such as saliency maps, Grad-CAM, SHAP, and counterfactuals often do not align with clinicians' reasoning processes and risk communication needs. The lack of standardised explainability benchmarks perpetuates legal ambiguity around liability and undermines clinician trust.
- **Systemic Algorithmic Bias and Equity Gaps:** Persistent underrepresentation of minority populations and structurally marginalised communities in training data contributes to unequal diagnostic performance across demographic subgroups. Despite increased awareness, regulatory mandates for fairness auditing, demographic validation, and post-deployment equity monitoring remain limited across most jurisdictions.
- **Fragmented and Reactive Governance Architecture:** Global governance frameworks remain fragmented, with significant variability in regulatory maturity, enforceability, and risk stratification approaches. Most existing models focus heavily on premarket approval while offering limited mechanisms for ongoing oversight, model drift detection, and adaptive regulatory updates. Australia, despite its innovation-friendly ethos, exemplifies this governance gap through the absence of comprehensive AI-specific healthcare legislation.
- **Underdeveloped Privacy-Preserving Infrastructure:** Although privacy-preserving approaches such as federated learning, homomorphic encryption, and secure multi-party computation offer promising pathways to ethical data sharing, their operational implementation remains rare. Institutional interoperability, technical complexity, and resource disparities continue to impede broader adoption in clinical practice.
- **Weak Stakeholder Engagement and Participatory Oversight:** The development and deployment of AI systems in healthcare remain largely siloed within technical, regulatory, or commercial domains, with insufficient participation from patients, clinicians, and marginalised communities. Stakeholder consultations often occur late in the design process or operate in advisory roles without decision-making authority, undermining the legitimacy and social robustness of AI governance structures.

5.2. Strategic Research Priorities for Responsible AI Governance

Addressing these limitations requires a coordinated and multidisciplinary research agenda that extends beyond technical optimisation to ethically embedded innovation. Key research priorities include:

- **Co-Governance Framework Development:** Future research should focus on empirically grounded governance models that integrate regulators, developers, clinicians, patients, ethicists, and civil society into ongoing governance processes. These frameworks should support adaptive regulation, continuous auditing, and sector-specific

accountability while balancing domestic legal requirements with international harmonisation efforts.

- **Clinically Relevant Fairness and Explainability Metrics:** Research should prioritise fairness metrics that extend beyond technical parity to incorporate clinical relevance, subgroup stratification, and context-sensitive outcome evaluation. Likewise, explainability tools should be assessed not only for transparency but also for their alignment with clinical reasoning and medicolegal requirements.
- **Institutionalisation of Participatory AI Design Practices:** Structured stakeholder inclusion methodologies should be developed to ensure that patients, marginalised groups, and frontline clinicians have meaningful input throughout the AI lifecycle, from problem formulation to post-market surveillance.
- **Ethical Evaluation of Privacy-Preserving AI Infrastructure:** Technical assessments of privacy-preserving technologies should be complemented by ethical evaluations examining real-world feasibility, institutional trustworthiness, and alignment with diverse data protection regimes.
- **Lifecycle Accountability and Post-Deployment Governance:** New governance mechanisms are needed to address post-market challenges, including ethical sunset clauses, adaptive decommissioning protocols, real-time auditing tools, and responsive regulatory triggers aligned with evolving system behaviour and clinical environments.

5.3. The Australian Context: A Normative Governance Imperative

For Australia, the ethical integration of AI into healthcare is at a regulatory crossroads. Although the Therapeutic Goods Administration (TGA) has demonstrated leadership through regulatory initiatives and voluntary ethics principles, the absence of enforceable AI-specific legislation remains a significant vulnerability as adaptive and generative AI systems enter clinical workflows. Continued reliance on institutional discretion, self-regulation, and fragmented pilot initiatives risks creating an uneven governance landscape characterised by inconsistent standards, limited post-market accountability, and unresolved liability concerns.

Australia's strong healthcare system, research capability, and early engagement with AI ethics position it to pioneer a globally relevant governance model. Such a model should combine the following:

- **Legally enforceable sector-specific AI legislation**, incorporating risk-tiered approval pathways, mandatory fairness audits, real-world evidence generation, and adaptive post-market surveillance.
- **Codified explainability and accountability standards**, which require human-interpretable model rationales that support both clinical safety and patient autonomy in informed consent processes.
- **Participatory and culturally responsive governance structures**, ensuring that indigenous populations, rural communities, and marginalised stakeholders are actively included in the governance of AI tools deployed in their healthcare contexts.
- **Global interoperability aligned with international best practices**, drawing selectively from leading models such as the EU AI Act 2024, FDA SaMD frameworks and WHO ethical principles, while tailoring implementation to Australia's unique sociotechnical ecosystem.

5.4. Toward a Governance Foundation for Ethical AI Innovation

Ultimately, trustworthy AI in healthcare is not solely a technical objective but also a governance and societal imperative. Without enforceable mechanisms that embed fairness,

transparency, and accountability throughout the AI lifecycle, public trust and clinical confidence may be undermined.

Sustainable AI integration in diagnostic imaging therefore depends on strengthening digital health governance infrastructures capable of coordinating innovation, regulation, and clinical accountability across interconnected healthcare systems.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/healthcare14131975/s1>, Supplementary File S1: PRISMA-ScR Reporting Checklist; Supplementary File S2: Database-Specific Search Strategies and Search Dates; Supplementary File S3: Included Studies and Thematic Classification Matrix (Excerpt).

Author Contributions: Conceptualization, D.A., K.A. and R.N.; Methodology, D.A.; Validation, D.A., K.A. and R.N.; Formal Analysis, D.A.; Investigation, D.A.; Resources, K.A. and R.N.; Data Curation, D.A.; Writing—Original Draft Preparation, D.A.; Writing—Review & Editing, D.A., K.A. and R.N.; Visualization, D.A.; Supervision, K.A. and R.N.; Project Administration, K.A. and R.N. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

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

Acknowledgments: The authors would like to thank Victoria University for providing access to institutional library resources and databases used in conducting this systematic survey.

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

References

1. Bi, W.L.; Hosny, A.; Schabath, M.B.; Giger, M.L.; Birkbak, N.J.; Mehrtash, A.; Allison, T.; Arnaout, O.; Abbosh, C.; Dunn, I.F.; et al. Artificial intelligence in cancer imaging: Clinical challenges and applications. *CA A Cancer J. Clin.* **2019**, *69*, 127–157. <https://doi.org/10.3322/caac.21552>.
2. Habli, I.; Lawton, T.; Porter, Z. Artificial intelligence in health care: Accountability and safety. *Bull. World Health Organ.* **2020**, *98*, 251. <https://doi.org/10.2471/BLT.19.237487>.
3. Guo, Y.; Hao, Z.; Zhao, S.; Gong, J.; Yang, F. Artificial intelligence in health care: Bibliometric analysis. *J. Med. Internet Res.* **2020**, *22*, e18228. <https://doi.org/10.2196/18228>.
4. Jimma, B.L. Artificial intelligence in healthcare: A bibliometric analysis. *Telemat. Inform. Rep.* **2023**, *9*, 100041. <https://doi.org/10.1016/j.teler.2023.100041>.
5. Chen, M.; Decary, M. Artificial intelligence in healthcare: An essential guide for health leaders. In *Healthcare Management Forum*; Sage Publications: Los Angeles, CA, USA, 2020; Volume 33, pp. 10–18. <https://doi.org/10.1177/0840470419873123>.
6. Kumar, P.; Chauhan, S.; Awasthi, L.K. Artificial intelligence in healthcare: Review, ethics, trust challenges & future research directions. *Eng. Appl. Artif. Intell.* **2023**, *120*, 105894. <https://doi.org/10.1016/j.engappai.2023.105894>.
7. Holmes, J.; Sacchi, L.; Bellazzi, R. Artificial intelligence in medicine. *Ann. R. Coll. Surg. Engl.* **2004**, *86*, 334–338.
8. Liu, P.; Lu, L.; Zhang, J.Y.; Huo, T.; Liu, S.X.; Ye, Z.W. Application of artificial intelligence in medicine: An overview. *Curr. Med. Sci.* **2021**, *41*, 1105–1115. <https://doi.org/10.1007/s11596-021-2474-3>.
9. Beam, A.L.; Drazen, J.M.; Kohane, I.S.; Leong, T.Y.; Manrai, A.K.; Rubin, E.J. Artificial intelligence in medicine. *N. Engl. J. Med.* **2023**, *388*, 1220–1221. <https://doi.org/10.1056/NEJMe2206291>.
10. Farhud, D.D.; Zokaei, S. Ethical issues of artificial intelligence in medicine and healthcare. *Iran. J. Public Health* **2021**, *50*, i–v. <https://doi.org/10.18502/ijph.v50i11.7600>.
11. Bekbolatova, M.; Mayer, J.; Ong, C.W.; Toma, M. Transformative Potential of AI in Healthcare: Definitions, Applications, and Navigating the Ethical Landscape and Public Perspectives. *Healthcare* **2024**, *12*, 125. <https://doi.org/10.3390/healthcare12020125>.
12. Weiner, E.B.; Dankwa-Mullan, I.; Nelson, W.A.; Hassanpour, S. Ethical challenges and evolving strategies in the integration of artificial intelligence into clinical practice. *PLoS Digit. Health* **2025**, *4*, e0000810. <https://doi.org/10.1371/journal.pdig.0000810>.
13. Najjar, R. Redefining radiology: A review of artificial intelligence integration in medical imaging. *Diagnostics* **2023**, *13*, 2760. <https://doi.org/10.3390/diagnostics13172760>.

14. Ahmad, Z.; Rahim, S.; Zubair, M.; Abdul-Ghafar, J. Artificial intelligence (AI) in medicine, current applications and future role with special emphasis on its potential and promise in pathology: present and future impact, obstacles including costs and acceptance among pathologists, practical and philosophical considerations. A comprehensive review. *Diagn. Pathol.* **2021**, *16*, 24. <https://doi.org/10.1186/s13000-021-01085-4>.
15. Barragán-Montero, A.; Javaid, U.; Valdés, G.; Nguyen, D.; Desbordes, P.; Macq, B.; Willems, S.; Vandewinckele, L.; Holmström, M.; Löfman, F.; et al. Artificial intelligence and machine learning for medical imaging: A technology review. *Phys. Medica* **2021**, *83*, 242–256. <https://doi.org/10.1016/j.ejmp.2021.04.016>.
16. Huynh, E.; Hosny, A.; Guthrie, C.; Bitterman, D.S.; Petit, S.F.; Haas-Kogan, D.A.; Kann, B.; Aerts, H.J.; Mak, R.H. Artificial intelligence in radiation oncology. *Nat. Rev. Clin. Oncol.* **2020**, *17*, 771–781. <https://doi.org/10.1038/s41571-020-0417-8>.
17. Thompson, R.F.; Valdes, G.; Fuller, C.D.; Carpenter, C.M.; Morin, O.; Aneja, S.; Lindsay, W.D.; Aerts, H.J.; Agrimson, B.; Deville, C., Jr.; et al. Artificial intelligence in radiation oncology: A specialty-wide disruptive transformation? *Radiother. Oncol.* **2018**, *129*, 421–426. <https://doi.org/10.1016/j.radonc.2018.05.030>.
18. Ahmed, Z.; Mohamed, K.; Zeeshan, S.; Dong, X. Artificial intelligence with multi-functional machine learning platform development for better healthcare and precision medicine. *Database* **2020**, *2020*, baaa010. <https://doi.org/10.1093/database/baaa010>.
19. Vaisman, A.; Linder, N.; Lundin, J.; Orchanian-Cheff, A.; Coulibaly, J.T.; Ephraim, R.K.; Bogoch, I.I. Artificial intelligence, diagnostic imaging and neglected tropical diseases: Ethical implications. *Bull. World Health Organ.* **2020**, *98*, 288. <https://doi.org/10.2471/BLT.19.237560>.
20. Petersson, L.; Larsson, I.; Nygren, J.M.; Nilsen, P.; Neher, M.; Reed, J.E.; Tyskbo, D.; Svedberg, P. Challenges to implementing artificial intelligence in healthcare: A qualitative interview study with healthcare leaders in Sweden. *BMC Health Serv. Res.* **2022**, *22*, 850. <https://doi.org/10.1186/s12913-022-08215-8>.
21. Iqbal, M.J.; Javed, Z.; Sadia, H.; Qureshi, I.A.; Irshad, A.; Ahmed, R.; Malik, K.; Raza, S.; Abbas, A.; Pezzani, R.; et al. Clinical applications of artificial intelligence and machine learning in cancer diagnosis: Looking into the future. *Cancer Cell Int.* **2021**, *21*, 270. <https://doi.org/10.1186/s12935-021-01981-1>.
22. Beil, M.; Proft, I.; Van Heerden, D.; Sviri, S.; Van Heerden, P.V. Ethical considerations about artificial intelligence for prognostication in intensive care. *Intensive Care Med. Exp.* **2019**, *7*, 70. <https://doi.org/10.1186/s40635-019-0286-6>.
23. McCradden, M.D.; Joshi, S.; Mazwi, M.; Anderson, J.A. Ethical limitations of algorithmic fairness solutions in health care machine learning. *Lancet Digit. Health* **2020**, *2*, e221–e223. [https://doi.org/10.1016/S2589-7500\(20\)30065-0](https://doi.org/10.1016/S2589-7500(20)30065-0).
24. Chen, I.Y.; Pierson, E.; Rose, S.; Joshi, S.; Ferryman, K.; Ghassemi, M. Ethical machine learning in healthcare. *Annu. Rev. Biomed. Data Sci.* **2021**, *4*, 123–144. <https://doi.org/10.1146/annurev-biodatasci-092820-114757>.
25. Rasheed, K.; Qayyum, A.; Ghaly, M.; Al-Fuqaha, A.; Razi, A.; Qadir, J. Explainable, trustworthy, and ethical machine learning for healthcare: A survey. *Comput. Biol. Med.* **2022**, *149*, 106043. <https://doi.org/10.1016/j.compbiomed.2022.106043>.
26. Albahri, A.S.; Duham, A.M.; Fadhel, M.A.; Alnoor, A.; Baqer, N.S.; Alzubaidi, L.; Albahri, O.S.; Alamoodi, A.H.; Bai, J.; Salhi, A.; et al. A systematic review of trustworthy and explainable artificial intelligence in healthcare: Assessment of quality, bias risk, and data fusion. *Inf. Fusion* **2023**, *96*, 156–191. <https://doi.org/10.1016/j.inffus.2023.03.012>.
27. Kaissis, G.A.; Makowski, M.R.; Rückert, D.; Braren, R.F. Secure, privacy-preserving and federated machine learning in medical imaging. *Nat. Mach. Intell.* **2020**, *2*, 305–311. <https://doi.org/10.1038/s42256-020-0186-1>.
28. Jobin, A.; Ienca, M.; Vayena, E. The global landscape of AI ethics guidelines. *Nat. Mach. Intell.* **2019**, *1*, 389–399. <https://doi.org/10.1038/s42256-019-0088-2>.
29. Shaheen, M.Y. Applications of Artificial Intelligence (AI) in healthcare: A review. *Sci. Prepr.* **2021**. <https://doi.org/10.14293/S2199-1006.1.SOR-PPLEUY.v1>.
30. Sun, Q.; Akman, A.; Schuller, B.W. Explainable artificial intelligence for medical applications: A review. *ACM Trans. Comput. Healthc.* **2025**, *6*, 1–31. <https://doi.org/10.1145/3701717>.
31. Char, D.S.; Abramoff, M.D.; Feudtner, C. Identifying ethical considerations for machine learning healthcare applications. *Am. J. Bioeth.* **2020**, *20*, 7–17. <https://doi.org/10.1080/15265161.2020.1819469>.
32. Zhang, J.; Zhang, Z.M. Ethics and governance of trustworthy medical artificial intelligence. *BMC Med. Inform. Decis. Mak.* **2023**, *23*, 7. <https://doi.org/10.1186/s12911-023-02103-9>.
33. Geis, J.R.; Brady, A.P.; Wu, C.C.; Spencer, J.; Ranschaert, E.; Jaremko, J.L.; Langer, S.G.; Borondy Kitts, A.; Birch, J.; Shields, W.F.; et al. Ethics of artificial intelligence in radiology: Summary of the joint European and North American multisociety statement. *Radiology* **2019**, *293*, 436–440. <https://doi.org/10.1148/radiol.2019191586>.
34. Amann, J.; Blasimme, A.; Vayena, E.; Frey, D.; Madai, V.I.; Consortium, P. Explainability for artificial intelligence in healthcare: A multidisciplinary perspective. *BMC Med. Inform. Decis. Mak.* **2020**, *20*, 310. <https://doi.org/10.1186/s12911-020-01332-6>.
35. Goirand, M.; Austin, E.; Clay-Williams, R. Implementing ethics in healthcare AI-based applications: A scoping review. *Sci. Eng. Ethics* **2021**, *27*, 61. <https://doi.org/10.1007/s11948-021-00336-3>.

36. Recht, M.P.; Dewey, M.; Dreyer, K.; Langlotz, C.; Niessen, W.; Prainsack, B.; Smith, J.J. Integrating artificial intelligence into the clinical practice of radiology: Challenges and recommendations. *Eur. Radiol.* **2020**, *30*, 3576–3584. <https://doi.org/10.1007/s00330-020-06672-5>.
37. Nichols, J.A.; Herbert Chan, H.W.; Baker, M.A. Machine learning: Applications of artificial intelligence to imaging and diagnosis. *Biophys. Rev.* **2019**, *11*, 111–118. <https://doi.org/10.1007/s12551-018-0449-9>.
38. Sun, T.Q.; Medaglia, R. Mapping the challenges of Artificial Intelligence in the public sector: Evidence from public healthcare. *Gov. Inf. Q.* **2019**, *36*, 368–383. <https://doi.org/10.1016/j.giq.2018.09.008>.
39. Grunhut, J.; Marques, O.; Wyatt, A.T. Needs, challenges, and applications of artificial intelligence in medical education curriculum. *JMIR Med. Educ.* **2022**, *8*, e35587. <https://doi.org/10.2196/35587>.
40. Reyes, M.; Meier, R.; Pereira, S.; Silva, C.A.; Dahlweid, F.M.; Tengg-Kobligk, H.v.; Summers, R.M.; Wiest, R. On the interpretability of artificial intelligence in radiology: challenges and opportunities. *Radiol. Artif. Intell.* **2020**, *2*, e190043. <https://doi.org/10.1148/ryai.2020190043>.
41. Chew, H.S.J.; Achananuparp, P. Perceptions and needs of artificial intelligence in health care to increase adoption: Scoping review. *J. Med. Internet Res.* **2022**, *24*, e32939. <https://doi.org/10.2196/32939>.
42. Lai, M.C.; Brian, M.; Mamzer, M.F. Perceptions of artificial intelligence in healthcare: Findings from a qualitative survey study among actors in France. *J. Transl. Med.* **2020**, *18*, 14. <https://doi.org/10.1186/s12967-019-02204-y>.
43. Olczak, J.; Pavlopoulos, J.; Prijs, J.; Ijpmma, F.F.; Doornberg, J.N.; Lundström, C.; Hedlund, J.; Gordon, M. Presenting artificial intelligence, deep learning, and machine learning studies to clinicians and healthcare stakeholders: An introductory reference with a guideline and a Clinical AI Research (CAIR) checklist proposal. *Acta Orthop.* **2021**, *92*, 513–525. <https://doi.org/10.1080/17453674.2021.1918389>.
44. Khalid, N.; Qayyum, A.; Bilal, M.; Al-Fuqaha, A.; Qadir, J. Privacy-preserving artificial intelligence in healthcare: Techniques and applications. *Comput. Biol. Med.* **2023**, *158*, 106848. <https://doi.org/10.1016/j.compbiomed.2023.106848>.
45. Larson, D.B.; Harvey, H.; Rubin, D.L.; Irani, N.; Tse, J.R.; Langlotz, C.P. Regulatory frameworks for development and evaluation of artificial intelligence-based diagnostic imaging algorithms: Summary and recommendations. *J. Am. Coll. Radiol.* **2021**, *18*, 413–424. <https://doi.org/10.1016/j.jacr.2020.09.060>.
46. Boudierhem, R. Shaping the future of AI in healthcare through ethics and governance. *Humanit. Soc. Sci. Commun.* **2024**, *11*, 416. <https://doi.org/10.1057/s41599-024-02894-w>.
47. Arnold, M. Teasing out artificial intelligence in medicine: An ethical critique of artificial intelligence and machine learning in medicine. *J. Bioethical Inq.* **2021**, *18*, 121–139. <https://doi.org/10.1007/s11673-020-10080-1>.
48. Morley, J.; Machado, C.C.; Burr, C.; Cows, J.; Joshi, I.; Taddeo, M.; Floridi, L. The ethics of AI in health care: A mapping review. *Soc. Sci. Med.* **2020**, *260*, 113172. <https://doi.org/10.1016/j.socscimed.2020.113172>.
49. Li, F.; Ruijs, N.; Lu, Y. Ethics & AI: A systematic review on ethical concerns and related strategies for designing with AI in healthcare. *AI* **2022**, *4*, 28–53. <https://doi.org/10.3390/ai4010002>.
50. Siala, H.; Wang, Y. SHIFTing artificial intelligence to be responsible in healthcare: A systematic review. *Soc. Sci. Med.* **2022**, *296*, 114782. <https://doi.org/10.1016/j.socscimed.2022.114782>.
51. Ahuja, A.S. The impact of artificial intelligence in medicine on the future role of the physician. *PeerJ* **2019**, *7*, e7702. <https://doi.org/10.7717/peerj.7702>.
52. Doraiswamy, P.M.; Blease, C.; Bodner, K. Artificial intelligence and the future of psychiatry: Insights from a global physician survey. *Artif. Intell. Med.* **2020**, *102*, 101753. <https://doi.org/10.1016/j.artmed.2019.101753>.
53. Krittanawong, C. The rise of artificial intelligence and the uncertain future for physicians. *Eur. J. Intern. Med.* **2018**, *48*, e13–e14. <https://doi.org/10.1016/j.ejim.2017.06.017>.
54. Yan, Y.; Zhang, J.W.; Zang, G.Y.; Pu, J. The primary use of artificial intelligence in cardiovascular diseases: What kind of potential role does artificial intelligence play in future medicine? *J. Geriatr. Cardiol. JGC* **2019**, *16*, 585. <https://doi.org/10.11909/j.issn.1671-5411.2019.08.010>.
55. Grunhut, J.; Wyatt, A.T.; Marques, O. Educating future physicians in artificial intelligence (AI): An integrative review and proposed changes. *J. Med. Educ. Curric. Dev.* **2021**, *8*, 23821205211036836. <https://doi.org/10.1177/23821205211036836>.
56. Zuhair, V.; Babar, A.; Ali, R.; Oduoye, M.O.; Noor, Z.; Chris, K.; Okon, I.I.; Rehman, L.U. Exploring the impact of artificial intelligence on global health and enhancing healthcare in developing nations. *J. Prim. Care Community Health* **2024**, *15*, 21501319241245847. <https://doi.org/10.1177/21501319241245847>.
57. Gala, D.; Behl, H.; Shah, M.; Makaryus, A.N. The role of artificial intelligence in improving patient outcomes and future of healthcare delivery in cardiology: A narrative review of the literature. *Healthcare* **2024**, *12*, 481. <https://doi.org/10.3390/healthcare12040481>.
58. Blease, C.; Worthen, A.; Torous, J. Psychiatrists' experiences and opinions of generative artificial intelligence in mental healthcare: An online mixed methods survey. *Psychiatry Res.* **2024**, *333*, 115724. <https://doi.org/10.1016/j.psychres.2024.115724>.
59. Grzybowski, A.; Jin, K.; Wu, H. Challenges of artificial intelligence in medicine and dermatology. *Clin. Dermatol.* **2024**, *42*, 210–215. <https://doi.org/10.1016/j.clindermatol.2023.12.017>.

60. Johnson, M.; Patel, M.; Phipps, A.; Van der Schaar, M.; Boulton, D.; Gibbs, M. The potential and pitfalls of artificial intelligence in clinical pharmacology. *CPT Pharmacometrics Syst. Pharmacol.* **2023**, *12*, 279. <https://doi.org/10.1002/psp4.12919>.
61. Aung, Y.Y.; Wong, D.C.; Ting, D.S. The promise of artificial intelligence: A review of the opportunities and challenges of artificial intelligence in healthcare. *Br. Med. Bull.* **2021**, *139*, 4–15. <https://doi.org/10.1093/bmb/ldab016>.
62. Wubineh, B.Z.; Deriba, F.G.; Woldeyohannis, M.M. Exploring the opportunities and challenges of implementing artificial intelligence in healthcare: A systematic literature review. In *Urologic Oncology: Seminars and Original Investigations*; Elsevier: Amsterdam, The Netherlands, 2024; Volume 42, pp. 48–56. <https://doi.org/10.1016/j.urolonc.2023.12.001>.
63. Olawade, D.B.; David-Olawade, A.C.; Wada, O.Z.; Asaolu, A.J.; Adereni, T.; Ling, J. Artificial intelligence in healthcare delivery: Prospects and pitfalls. *J. Med. Surg. Public Health* **2024**, *3*, 100108. <https://doi.org/10.1016/j.glmedi.2024.100108>.
64. Bhagat, S.V.; Kanyal, D. Navigating the future: The transformative impact of artificial intelligence on hospital management—A comprehensive review. *Cureus* **2024**, *16*, e54518. <https://doi.org/10.7759/cureus.54321>.
65. Bellini, V.; Russo, M.; Domenichetti, T.; Panizzi, M.; Allai, S.; Bignami, E.G. Artificial intelligence in operating room management. *J. Med. Syst.* **2024**, *48*, 19. <https://doi.org/10.1007/s10916-024-02038-2>.
66. Khang, A. *AI and IoT Technology and Applications for Smart Healthcare Systems*; CRC Press: Boca Raton, FL, USA, 2024.
67. Chen, X.; Xie, H.; Tao, X.; Wang, F.L.; Leng, M.; Lei, B. Artificial intelligence and multimodal data fusion for smart healthcare: Topic modeling and bibliometrics. *Artif. Intell. Rev.* **2024**, *57*, 91. <https://doi.org/10.1007/s10462-024-10712-7>.
68. Pradyumna, G.; Hegde, R.B.; Bommegowda, K.; Jan, T.; Naik, G.R. Empowering healthcare with IoMT: Evolution, machine learning integration, security, and interoperability challenges. *IEEE Access* **2024**, *12*, 20603–20623. <https://doi.org/10.1109/ACCESS.2024.3362239>.
69. Chen, R.J.; Wang, J.J.; Williamson, D.F.; Chen, T.Y.; Lipkova, J.; Lu, M.Y.; Sahai, S.; Mahmood, F. Algorithmic fairness in artificial intelligence for medicine and healthcare. *Nat. Biomed. Eng.* **2023**, *7*, 719–742. <https://doi.org/10.1038/s41551-023-01056-8>.
70. Mennella, C.; Maniscalco, U.; De Pietro, G.; Esposito, M. Ethical and regulatory challenges of AI technologies in healthcare: A narrative review. *Heliyon* **2024**, *10*, e26297. <https://doi.org/10.1016/j.heliyon.2024.e26176>.
71. Messina, P.; Pino, P.; Parra, D.; Soto, A.; Besa, C.; Uribe, S.; Andía, M.; Tejos, C.; Prieto, C.; Capurro, D. A survey on deep learning and explainability for automatic report generation from medical images. *ACM Comput. Surv.* **2022**, *54*, 1–40. <https://doi.org/10.1145/3522747>.
72. Paproki, A.; Salvado, O.; Fookes, C. Synthetic data for deep learning in computer vision & medical imaging: A means to reduce data bias. *ACM Comput. Surv.* **2024**, *56*, 1–37. <https://doi.org/10.1145/3648570>.
73. Hossain, M.I.; Zamzmi, G.; Mouton, P.R.; Salekin, M.S.; Sun, Y.; Goldgof, D. Explainable AI for medical data: Current methods, limitations, and future directions. *ACM Comput. Surv.* **2025**, *57*, 1–46. <https://doi.org/10.1145/3637487>.
74. Amann, J.; Bürger, V.K.; Livne, M.; Bui, C.K.; Madai, V.I. The fundamentals of AI ethics in medical imaging. In *Trustworthy AI in Medical Imaging*; Elsevier: Amsterdam, The Netherlands, 2025; pp. 7–33. <https://doi.org/10.1016/B978-0-44-323761-4.00010-9>.
75. Jordan, M.I.; Mitchell, T.M. Machine learning: Trends, perspectives, and prospects. *Science* **2015**, *349*, 255–260. <https://doi.org/10.1126/science.aaa8415>.
76. Goisauf, M.; Cano Abadía, M. Ethics of AI in radiology: A review of ethical and societal implications. *Front. Big Data* **2022**, *5*, 850383. <https://doi.org/10.3389/fdata.2022.850383>.
77. Herington, J.; McCradden, M.D.; Creel, K.; Boellaard, R.; Jones, E.C.; Jha, A.K.; Rahmim, A.; Scott, P.J.; Sunderland, J.J.; Wahl, R.L.; et al. Ethical considerations for artificial intelligence in medical imaging: Deployment and governance. *J. Nucl. Med.* **2023**, *64*, 1509–1515. <https://doi.org/10.2967/jnumed.123.266110>.
78. Jiang, L.; Wu, Z.; Xu, X.; Zhan, Y.; Jin, X.; Wang, L.; Qiu, Y. Opportunities and challenges of artificial intelligence in the medical field: Current application, emerging problems, and problem-solving strategies. *J. Int. Med. Res.* **2021**, *49*, 03000605211000157. <https://doi.org/10.1177/03000605211000157>.
79. Theriault-Lauzier, P.; Cobin, D.; Tastet, O.; Langlais, E.L.; Taji, B.; Kang, G.; Chong, A.Y.; So, D.; Tang, A.; Gichoya, J.W.; et al. A responsible framework for applying artificial intelligence on medical images and signals at the point-of-care: The PACS-AI platform. *Can. J. Cardiol.* **2024**. <https://doi.org/10.1016/j.cjca.2024.01.025>.
80. Kulkov, I. Next-generation business models for artificial intelligence start-ups in the healthcare industry. *Int. J. Entrep. Behav. Res.* **2023**, *29*, 860–885. <https://doi.org/10.1108/IJEBR-04-2021-0304>.
81. Chikhaoui, E.; Alajmi, A.; Larabi-Marie-Sainte, S. Artificial intelligence applications in healthcare sector: Ethical and legal challenges. *Emerg. Sci. J.* **2022**, *6*, 717–738. <https://doi.org/10.28991/ESJ-2022-06-04-08>.
82. Saraswat, D.; Bhattacharya, P.; Verma, A.; Prasad, V.K.; Tanwar, S.; Sharma, G.; Bokoro, P.N.; Sharma, R. Explainable AI for healthcare 5.0: Opportunities and challenges. *IEEE Access* **2022**, *10*, 84486–84517. <https://doi.org/10.1109/ACCESS.2022.3197671>.
83. Aminizadeh, S.; Heidari, A.; Dehghan, M.; Toumaj, S.; Rezaei, M.; Navimipour, N.J.; Stroppa, F.; Unal, M. Opportunities and challenges of artificial intelligence and distributed systems to improve the quality of healthcare service. *Artif. Intell. Med.* **2024**, *149*, 102779. <https://doi.org/10.1016/j.artmed.2024.102779>.

84. Tulgar, Y.K.; Tulgar, S.; Köse, S.G.; Köse, H.C.; Nasirlier, G.Ç.; Doğan, M.; Thomas, D.T. Anesthesiologists' perspective on the use of artificial intelligence in ultrasound-guided regional anaesthesia in terms of medical ethics and medical education: A survey study. *Eurasian J. Med.* **2023**, *55*, 146. <https://doi.org/10.5152/eurasianjmed.2023.22254>.
85. Karimian, G.; Petelos, E.; Evers, S.M. The ethical issues of the application of artificial intelligence in healthcare: A systematic scoping review. *AI Ethics* **2022**, *2*, 539–551. <https://doi.org/10.1007/s43681-021-00131-7>.
86. Nasir, S.; Khan, R.A.; Bai, S. Ethical framework for harnessing the power of AI in healthcare and beyond. *IEEE Access* **2024**, *12*, 31014–31035. <https://doi.org/10.1109/ACCESS.2024.3369912>.
87. Zahlan, A.; Ranjan, R.P.; Hayes, D. Artificial intelligence innovation in healthcare: Literature review, exploratory analysis, and future research. *Technol. Soc.* **2023**, *74*, 102321. <https://doi.org/10.1016/j.techsoc.2023.102321>.
88. Zangana, H.M.; Sallow, Z.B.; Salih, B.A. The Impact of Artificial Intelligence on Healthcare: A Systematic Review of Innovations, Challenges, and Ethical Considerations. *J. Comput. Digit. Bus.* **2025**, *4*, 1–9. <https://doi.org/10.56427/jcbd.v4i1.601>.
89. Almeida-Galárraga, D.; Tirado-Espín, A. Acceptability of AI in Medical Diagnostics: A Discourse Analysis. *Commun. Appl. Technol. Proc. ICOMTA 2024* **2025**, *427*, 383. https://doi.org/10.1007/978-981-96-0426-5_33.
90. Nowrozy, R.; Ahmed, K.; Wang, H. GPT, ontology, and CAABAC: A tripartite personalized access control model anchored by compliance, context and attribute. *PLoS ONE* **2025**, *20*, e0310553. <https://doi.org/10.1371/journal.pone.0310553>.
91. Nowrozy, R.; Ahmed, K.; Wang, H.; McIntosh, T. Towards a universal privacy model for electronic health record systems: An ontology and machine learning approach. *Informatics* **2023**, *10*, 60. <https://doi.org/10.3390/informatics10030060>.
92. Nowrozy, R. A Security and Privacy Compliant Data Sharing Solution For Healthcare Data Ecosystems. Ph.D. Thesis, Victoria University, Melbourne, VIC, Australia, 2024.
93. Shrotriya, V.; Jain, M. A.; Shrivastava, P.; Sharma, P. Exploring IoT solutions for connecting and synchronizing various maternal health monitoring devices. In *Interdisciplinary Work of Science and Technology in Maternal and Child Care*; Xoffencer International Book Publication House: Gwalior, India, 2019; p. 100.
94. Ntjamba, F.C.; Ashipala, D.O. Impact on and Ethical Considerations of Artificial Intelligence on Human Healthcare. In *AI Technologies and Advancements for Psychological Well-Being and Healthcare*; IGI Global: IGI Global, PA, USA, 2025; pp. 1–36. <https://doi.org/10.4018/979-8-3693-9158-7.ch001>.
95. Pantelopoulos, A.; Bourbakis, N.G. A survey on wearable sensor-based systems for health monitoring and prognosis. *IEEE Trans. Syst. Man Cybern. Part C (Appl. Rev.)* **2009**, *40*, 1–12. <https://doi.org/10.1109/TSMCC.2009.2032660>.
96. AlAmir, M.; AlGhamdi, M. The Role of generative adversarial network in medical image analysis: An in-depth survey. *ACM Comput. Surv.* **2022**, *55*, 1–36. <https://doi.org/10.1145/3527849>.
97. Stogiannos, N.; Georgiadou, E.; Rarri, N.; Malamateniou, C. Ethical AI: A qualitative study exploring ethical challenges and solutions on the use of AI in medical imaging. *Eur. J. Radiol. Artif. Intell.* **2025**, *1*, 100006. <https://doi.org/10.1016/j.ejrai.2025.100006>.
98. Wang, J.; Wang, K.; Yu, Y.; Lu, Y.; Xiao, W.; Sun, Z.; Liu, F.; Zou, Z.; Gao, Y.; Yang, L.; et al. Self-improving generative foundation model for synthetic medical image generation and clinical applications. *Nat. Med.* **2025**, *31*, 609–617. <https://doi.org/10.1038/s41591-024-03359-y>.
99. Koçak, B.; Ponsiglione, A.; Stanzione, A.; Bluethgen, C.; Santinha, J.; Ugga, L.; Huisman, M.; Klontzas, M.E.; Cannella, R.; Cuocolo, R. Bias in artificial intelligence for medical imaging: Fundamentals, detection, avoidance, mitigation, challenges, ethics, and prospects. *Diagn. Interv. Radiol.* **2025**, *31*, 75. <https://doi.org/10.4274/dir.2024.242854>.
100. Sand, M.; Durán, J.M.; Jongsma, K.R. Responsibility beyond design: Physicians' requirements for ethical medical AI. *Bioethics* **2022**, *36*, 162–169. <https://doi.org/10.1111/bioe.12963>.
101. Nowrozy, R.; Ahmed, K. A Systematic Survey on AI Governance in Healthcare. *ACM Comput. Surv.* **2024**, *56*, 1–41.
102. Megerian, J.T.; Dey, S.; Melmed, R.D.; Coury, D.L.; Lerner, M.; Nicholls, C.J.; Sohl, K.; Rouhbakhsh, R.; Narasimhan, A.; Romain, J.; et al. Evaluation of an artificial intelligence-based medical device for diagnosis of autism spectrum disorder. *NPJ Digit. Med.* **2022**, *5*, 57. <https://doi.org/10.1038/s41746-022-00604-x>.
103. Reddy, S.; Allan, S.; Coghlan, S.; Cooper, P. A governance model for the application of AI in health care. *J. Am. Med. Inform. Assoc.* **2020**, *27*, 491–497. <https://doi.org/10.1093/jamia/ocz192>.
104. Abrámoff, M.D.; Roehrenbeck, C.; Trujillo, S.; Goldstein, J.; Graves, A.S.; Repka, M.X.; Silva, Ezequiel Zeke, I. A reimbursement framework for artificial intelligence in healthcare. *NPJ Digit. Med.* **2022**, *5*, 72. <https://doi.org/10.1038/s41746-022-00621-w>.
105. Amjad, A.; Kordel, P.; Fernandes, G. A review on innovation in healthcare sector (telehealth) through artificial intelligence. *Sustainability* **2023**, *15*, 6655. <https://doi.org/10.3390/su15086655>.
106. El-Sherif, D.M.; Abouzid, M.; Elzarif, M.T.; Ahmed, A.A.; Albakri, A.; Alshehri, M.M. Telehealth and artificial intelligence insights into healthcare during the COVID-19 pandemic. *Healthcare* **2022**, *10*, 385. <https://doi.org/10.3390/healthcare10020385>.
107. Kotter, E.; D'Antonoli, T.A.; Cuocolo, R.; Hierath, M.; Huisman, M.; Klontzas, M.E.; Martí-Bonmatí, L.; May, M.S.; Neri, E.; Nikolaou, K.; et al. Guiding AI in radiology: ESR's recommendations for effective implementation of the European AI Act. *Insights Into Imaging* **2025**, *16*, 33. <https://doi.org/10.1186/s13244-025-01905-x>.

108. Corfmat, M.; Martineau, J.T.; Régis, C. High-reward, high-risk technologies? An ethical and legal account of AI development in healthcare. *BMC Med. Ethics* **2025**, *26*, 4. <https://doi.org/10.1186/s12910-024-01158-1>.
109. Hickman, S.E.; Baxter, G.C.; Gilbert, F.J. Adoption of artificial intelligence in breast imaging: Evaluation, ethical constraints and limitations. *Br. J. Cancer* **2021**, *125*, 15–22. <https://doi.org/10.1038/s41416-021-01333-w>.
110. Dave, T.; Athaluri, S.A.; Singh, S. ChatGPT in medicine: An overview of its applications, advantages, limitations, future prospects, and ethical considerations. *Front. Artif. Intell.* **2023**, *6*, 1169595. <https://doi.org/10.3389/frai.2023.1169595>.
111. Kothinti, R.R. Artificial intelligence in healthcare: Revolutionizing precision medicine, predictive analytics, and ethical considerations in autonomous diagnostics. *World J. Adv. Res. Rev.* **2024**, *19*, 3395–3406. <https://doi.org/10.30574/wjarr.2024.24.3.3675>.
112. Maleki Varnosfaderani, S.; Forouzanfar, M. The role of AI in hospitals and clinics: Transforming healthcare in the 21st century. *Bioengineering* **2024**, *11*, 337. <https://doi.org/10.3390/bioengineering11040337>.
113. Mörch, C.; Atsu, S.; Cai, W.; Li, X.; Madathil, S.; Liu, X.; Mai, V.; Tamimi, F.; Dilhac, M.; Ducret, M. Artificial intelligence and ethics in dentistry: A scoping review. *J. Dent. Res.* **2021**, *100*, 1452–1460. <https://doi.org/10.1177/00220345211013808>.
114. Al-kfairy, M.; Mustafa, D.; Kshetri, N.; Insiew, M.; Alfandi, O. Ethical challenges and solutions of generative AI: An interdisciplinary perspective. *Informatics* **2024**, *11*, 58. <https://doi.org/10.3390/informatics11030058>.
115. Čartolovni, A.; Tomičić, A.; Mosler, E.L. Ethical, legal, and social considerations of AI-based medical decision-support tools: A scoping review. *Int. J. Med. Inform.* **2022**, *161*, 104738. <https://doi.org/10.1016/j.ijmedinf.2022.104738>.
116. Dhar, T.; Dey, N.; Borra, S.; Sherratt, R.S. Challenges of deep learning in medical image analysis—Improving explainability and trust. *IEEE Trans. Technol. Soc.* **2023**, *4*, 68–75. <https://doi.org/10.1109/TTS.2023.3234203>.
117. Durán, J.M.; Jongsma, K.R. Who is afraid of black box algorithms? On the epistemological and ethical basis of trust in medical AI. *J. Med. Ethics* **2021**, *47*, 329–335. <https://doi.org/10.1136/medethics-2020-106820>.
118. Ullah, S.; Li, J.; Chen, J.; Ali, I.; Khan, S.; Ahad, A.; Ullah, F.; Leung, V.C. A Survey on Emerging Trends and Applications of 5G and 6G to Healthcare Environments. *ACM Comput. Surv.* **2024**, *57*, 1–36. <https://doi.org/10.1145/3699715>.
119. Hertel, R.; Benlamri, R. Deep learning techniques for COVID-19 diagnosis and prognosis based on radiological imaging. *ACM Comput. Surv.* **2023**, *55*, 1–39. <https://doi.org/10.1145/3573204>.
120. Wood, A.; Najarian, K.; Kahrobaei, D. Homomorphic encryption for machine learning in medicine and bioinformatics. *ACM Comput. Surv.* **2020**, *53*, 1–35. <https://doi.org/10.1145/3394658>.
121. Paladugu, P.S.; Ong, J.; Nelson, N.; Kamran, S.A.; Waisberg, E.; Zaman, N.; Kumar, R.; Dias, R.D.; Lee, A.G.; Tavakkoli, A. Generative adversarial networks in medicine: Important considerations for this emerging innovation in artificial intelligence. *Ann. Biomed. Eng.* **2023**, *51*, 2130–2142. <https://doi.org/10.1007/s10439-023-03304-z>.
122. Showrov, A.A.; Aziz, M.T.; Nabil, H.R.; Jim, J.R.; Kabir, M.M.; Mridha, M.; Asai, N.; Shin, J. Generative adversarial networks (GANs) in medical imaging: advancements, applications and challenges. *IEEE Access* **2024**, *12*, 35728–35753. <https://doi.org/10.1109/ACCESS.2024.3370848>.
123. Shafik, W. Generative Adversarial Networks: Security, Privacy, and Ethical Considerations. In *Generative Artificial Intelligence (AI) Approaches for Industrial Applications*; Springer: Berlin/Heidelberg, Germany, 2025; pp. 93–117.
124. Ranschaert, E.R.; Morozov, S.; Algra, P.R. *Artificial Intelligence in Medical Imaging: Opportunities, Applications and Risks*; Springer: Berlin/Heidelberg, Germany, 2019.
125. Koohi-Moghadam, M.; Bae, K.T. Generative AI in medical imaging: Applications, challenges, and ethics. *J. Med. Syst.* **2023**, *47*, 94. <https://doi.org/10.1007/s10916-023-01987-4>.
126. Nazar, M.; Alam, M.M.; Yafi, E.; Su’ud, M.M. A systematic review of human–computer interaction and explainable artificial intelligence in healthcare with artificial intelligence techniques. *IEEE Access* **2021**, *9*, 153316–153348. <https://doi.org/10.1109/ACCESS.2021.3127881>.
127. Giordano, C.; Brennan, M.; Mohamed, B.; Rashidi, P.; Modave, F.; Tighe, P. Accessing artificial intelligence for clinical decision-making. *Front. Digit. Health* **2021**, *3*, 645232. <https://doi.org/10.3389/fdgth.2021.645232>.
128. Bartoletti, I. AI in healthcare: Ethical and privacy challenges. In *Proceedings of the Artificial Intelligence in Medicine: 17th Conference on Artificial Intelligence in Medicine, AIME 2019*; Springer International Publishing: Cham, Switzerland, 2019; pp. 7–10. https://doi.org/10.1007/978-3-030-21642-9_2.
129. Blease, C.; Kaptchuk, T.J.; Bernstein, M.H.; Mandl, K.D.; Halamka, J.D.; DesRoches, C.M. Artificial intelligence and the future of primary care: Exploratory qualitative study of UK general practitioners’ views. *J. Med. Internet Res.* **2019**, *21*, e12802. <https://doi.org/10.2196/12802>.
130. Murphy, K.; Di Ruggiero, E.; Upshur, R.; Willison, D.J.; Malhotra, N.; Cai, J.C.; Malhotra, N.; Lui, V.; Gibson, J. Artificial intelligence for good health: A scoping review of the ethics literature. *BMC Med. Ethics* **2021**, *22*, 14. <https://doi.org/10.1186/s12910-021-00577-8>.
131. Lee, E.E.; Torous, J.; De Choudhury, M.; Depp, C.A.; Graham, S.A.; Kim, H.C.; Paulus, M.P.; Krystal, J.H.; Jeste, D.V. Artificial intelligence for mental health care: Clinical applications, barriers, facilitators, and artificial wisdom. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* **2021**, *6*, 856–864. <https://doi.org/10.1016/j.bpsc.2021.05.001>.

132. Huang, S.; Yang, J.; Fong, S.; Zhao, Q. Artificial intelligence in cancer diagnosis and prognosis: Opportunities and challenges. *Cancer Lett.* **2020**, *471*, 61–71. <https://doi.org/10.1016/j.canlet.2019.12.007>.
133. Mishra, R.; Satpathy, R.; Pati, B. Interpretable AI in Medical Imaging: Enhancing Diagnostic Accuracy through Human–Computer Interaction. *J. Artif. Intell. Syst.* **2024**, *6*, 96–111. <https://doi.org/10.33969/AIS.2024060107>.
134. Patrício, C.; Neves, J.C.; Teixeira, L.F. Explainable deep learning methods in medical image classification: A survey. *ACM Comput. Surv.* **2023**, *56*, 1–41. <https://doi.org/10.48550/arXiv.2205.04766>.
135. Jeyaraman, M.; Balaji, S.; Jeyaraman, N.; Yadav, S. Unraveling the ethical enigma: Artificial intelligence in healthcare. *Cureus* **2023**, *15*, e43262. <https://doi.org/10.7759/cureus.43900>.
136. Zhu, W.; Huang, L.; Zhou, X.; Li, X.; Shi, G.; Ying, J.; Wang, C. Could AI ethical anxiety, perceived ethical risks and ethical awareness about AI influence university students' use of generative AI products? An ethical perspective. *Int. J. Hum.-Comput. Interact.* **2025**, *41*, 742–764. <https://doi.org/10.1080/10447318.2024.2323277>.
137. Müller, H.; Mayrhofer, M.T.; Van Veen, E.B.; Holzinger, A. The Ten Commandments of Ethical Medical AI. *Computer* **2021**, *54*, 119–123. <https://doi.org/10.1109/MC.2021.3074263>.
138. Borenstein, J.; Howard, A. Emerging challenges in AI and the need for AI ethics education. *AI Ethics* **2021**, *1*, 61–65. <https://doi.org/10.1007/s43681-020-00002-7>.
139. Odle, T. The AI era: The role of medical imaging and radiation therapy professionals. *Radiol. Technol.* **2020**, *91*, 391–400.
140. Li, J.; Zhu, G.; Hua, C.; Feng, M.; Bennamoun, B.; Li, P.; Lu, X.; Song, J.; Shen, P.; Xu, X.; et al. A systematic collection of medical image datasets for deep learning. *ACM Comput. Surv.* **2023**, *56*, 1–51. <https://doi.org/10.1145/3615862>.
141. Dong, J.; Chen, J.; Xie, X.; Lai, J.; Chen, H. Survey on Adversarial Attack and Defense for Medical Image Analysis: Methods and Challenges. *ACM Comput. Surv.* **2024**, *57*, 1–38. <https://doi.org/10.1145/3702638>.
142. Thieme, A.; Rajamohan, A.; Cooper, B.; Groombridge, H.; Simister, R.; Wong, B.; Woznitza, N.; Pinnock, M.A.; Wetscherek, M.T.; Morrison, C.; et al. Challenges for Responsible AI Design and Workflow Integration in Healthcare: A Case Study of Automatic Feeding Tube Qualification in Radiology. *ACM Trans. Comput.-Hum. Interact.* **2025**, *32*, 1–36. <https://doi.org/10.1145/3716500>.
143. Chauhan, C.; Gullapalli, R.R. Ethics of AI in pathology: Current paradigms and emerging issues. In *Artificial Intelligence in Pathology*; Elsevier: Amsterdam, The Netherlands, 2025; pp. 159–180. <https://doi.org/10.1016/j.ajpath.2021.06.011>.
144. D'Antonoli, T.A. Ethical considerations for artificial intelligence: An overview of the current radiology landscape. *Diagn. Interv. Radiol.* **2020**, *26*, 504. <https://doi.org/10.5152/dir.2020.19279>.
145. Naik, N.; Hameed, B.; Shetty, D.K.; Swain, D.; Shah, M.; Paul, R.; Aggarwal, K.; Ibrahim, S.; Patil, V.; Smriti, K.; et al. Legal and ethical consideration in artificial intelligence in healthcare: Who takes responsibility? *Front. Surg.* **2022**, *9*, 862322. <https://doi.org/10.3389/fsurg.2022.862322>.
146. Pesapane, F.; Hauglid, M.K.; Fumagalli, M.; Petersson, L.; Parkar, A.P.; Cassano, E.; Horgan, D. The translation of in-house imaging AI research into a medical device ensuring ethical and regulatory integrity. *Eur. J. Radiol.* **2025**, *182*, 111852. <https://doi.org/10.1016/j.ejrad.2024.111852>.
147. Wang, P.; Gao, Q.; Wu, X.; You, T.; Du, H.; Wang, X.; Zeng, S.; Lv, Q.; Ding, X.; Wang, L.; et al. Ethics and Safety in Medical Imaging and Artificial Intelligence. In *Artificial Intelligence in Medical Imaging in China*; Springer: Berlin/Heidelberg, Germany, 2024; pp. 431–446. https://doi.org/10.1007/978-981-99-8441-1_29.
148. Currie, G.; Hawk, K.E.; Rohren, E.M. Ethical principles for the application of artificial intelligence (AI) in nuclear medicine. *Eur. J. Nucl. Med. Mol. Imaging* **2020**, *47*, 748–752. <https://doi.org/10.1007/s00259-020-04678-1>.
149. Arkoh, S.; Akudjedu, T.N.; Amedu, C.; Antwi, W.K.; Elshami, W.; Ohene-Botwe, B. Current Radiology workforce perspective on the integration of artificial intelligence in clinical practice: A systematic review. *J. Med. Imaging Radiat. Sci.* **2025**, *56*, 101769. <https://doi.org/10.1016/j.jmir.2024.101769>.
150. Priyadarshi, R.; Ranjan, R.; Vishwakarma, A.K.; Yang, T.; Rathore, R.S. Exploring the Frontiers of Unsupervised Learning Techniques for Diagnosis of Cardiovascular Disorder: A Systematic Review. *IEEE Access* **2024**, *12*, 139253–139272. <https://doi.org/10.1109/ACCESS.2024.3468163>.
151. Ganesan, S.; Somasiri, N. Navigating the Integration of Machine Learning in Healthcare: Challenges, Strategies, and Ethical Considerations. *J. Comput. Cogn. Eng.* **2024**, *4*, 8–23. <https://doi.org/10.47852/bonviewJCCE42023600>.
152. Nanjundan, P.; Indu, P.; Thomas, L. Navigating the ethical landscape of artificial intelligence: Challenges, frameworks, and responsible deployment. In *Artificial Intelligence Technologies for Engineering Applications*; CRC Press: Boca Raton, FL, USA, 2025; pp. 1–13. <https://doi.org/10.12785/ijcds/1571038843>.
153. Ji, S.; Li, X.; Sun, W.; Dong, H.; Taalas, A.; Zhang, Y.; Wu, H.; Pitkänen, E.; Marttinen, P. A unified review of deep learning for automated medical coding. *ACM Comput. Surv.* **2024**, *56*, 1–41. <https://doi.org/10.1145/3664615>.
154. Mohsin Khan, M.; Shah, N.; Shaikh, N.; Thabet, A.; alrabayah, T.; Belkhair, S. Towards secure and trusted AI in healthcare: A systematic review of emerging innovations and ethical challenges. *Int. J. Med. Inform.* **2025**, *195*, 105780. <https://doi.org/10.1016/j.ijmedinf.2024.105780>.

155. Saw, S.N.; Ng, K.H. Current challenges of implementing artificial intelligence in medical imaging. *Phys. Medica* **2022**, *100*, 12–17. <https://doi.org/10.1016/j.ejmp.2022.06.002>.
156. Aldhafeeri, F.M. Governing Artificial Intelligence in Radiology: A Systematic Review of Ethical, Legal, and Regulatory Frameworks. *Diagnostics* **2025**, *15*, 2300. <https://doi.org/10.3390/diagnostics15182300>.

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