UPM Innovation

Research think tank

https://upminnovation.org/

Artificial intelligence and robotics in cancer diagnostic imaging

From algorithmic performance to image-guided diagnostic infrastructure

Kinda Chebib[1] **Mickael Tardy**[2]

[1] UPM Innovation, Université Pour la Méditerranée, France

[2] Hera-Mi, Saint Herblain, France

Highlights

- AI can improve cancer diagnosis by stabilising interpretation and reducing diagnostic delays.
- Robotics can improve cancer diagnosis by stabilising biopsy targeting precision and sampling reproducibility.
- AI and robotics together close the diagnostic loop: detect → target → confirm.
- Micro-/nanorobotics is a credible emerging diagnostic frontier but remains largely translational.
- Sustainable adoption depends on integration, governance, and continuous performance oversight.

Abstract

Background

Cancer diagnosis increasingly depends on imaging-centred pathways in which radiology, image-guided intervention, and workflow governance jointly shape diagnostic performance. Artificial intelligence (AI) has demonstrated strong performance in cancer imaging tasks including lesion detection, risk stratification, and triage. In parallel, robotics has advanced image-guided diagnostic procedures, particularly biopsies, improving targeting precision and procedural reproducibility. Despite rapid technological progress, real-world clinical adoption of both remains uneven.

Objective

To synthesise clinical and translational evidence on how AI and robotics, independently and in combination, may improve cancer diagnosis in imaging, with a focus on diagnostic performance, workflow integration, governance, and system-level adoption beyond algorithmic accuracy alone.

Methods

This paper is an integrative narrative review supported by a structured scoping search of MEDLINE/PubMed, Scopus, Web of Science Core Collection, and IEEE Xplore (2018–2026), together with citation chasing and reproducible public-evidence triangulation to characterise adoption context. Evidence was extracted into an evidence map covering diagnostic step, workflow role, and governance attributes. Findings were interpreted using domain-appropriate appraisal lenses, including QUADAS-2/PROBAST and CLAIM/CONSORT-AI. No meta-analysis was performed. Adoption context for Europe and Saudi Arabia was assessed using public regulatory and institutional sources, and a brief illustrative screening-workflow example (Breast-SlimView) was grounded in peer-reviewed and publicly available documentation.

Results

AI appears to improve cancer diagnosis primarily by stabilising interpretation, reducing cognitive variability and accelerating time-to-review, although performance can degrade under dataset

shift and workflow misalignment. Robotics appears to improve diagnosis primarily by stabilising the confirmation step, reducing targeting variability and improving reproducibility in image-guided biopsy and minimally invasive sampling pathways. Illustrative findings include reduced radiologist workload under AI-supported mammography screening configurations and reduced needle adjustment or improved placement control in image-guided robotic confirmation workflows, although effect sizes remain context-dependent. Where this review refers to "signals" such as adoption, infrastructure, or governance signals, these are treated as contextual indicators of deployment readiness and not as evidence of clinical effectiveness; clinical-effectiveness claims are grounded only in peer-reviewed outcome data. Emerging micro- and nanorobotics suggest future diagnostic modes such as in situ sensing and targeted micro-scale interaction, but translation remains early and depends on progress in biocompatibility, controllability, tracking, and regulatory readiness.

Conclusion

AI and robotics function as complementary components of contemporary cancer diagnosis, addressing distinct but interdependent sources of diagnostic uncertainty. By distinguishing interpretive stabilisation from confirmatory stabilisation, this review reframes diagnostic value away from isolated performance metrics and toward variance reduction at the system level. Durable improvement in cancer diagnosis is most likely when AI and robotics are integrated into governed, end-to-end diagnostic infrastructure rather than deployed as standalone tools.

Keywords

Artificial intelligence; diagnostic imaging; robotics; cancer diagnosis; image-guided biopsy; micro-/nanorobots; governance

1. Introduction

Cancer diagnostic imaging is increasingly shaped by artificial intelligence (AI) and robotics. AI systems can detect lesions, quantify imaging features, and prioritise cases, supporting radiologists under growing workforce and throughput pressure [5–9]. Robotics, by contrast, can stabilise and standardise procedural steps in image-guided diagnosis, especially where targeting accuracy and sampling reproducibility determine diagnostic certainty [31,33–37]. Despite rapid progress, the translation of both AI and robotics into routine diagnostic infrastructure remains uneven. Many studies demonstrate strong algorithmic performance under controlled conditions, yet real-world implementation is influenced by workflow integration, governance, regulatory constraints, and health-system readiness [10,11,20].

In cancer imaging, uncertainty arises both during interpretation (reading/triage) and during confirmation (targeting and sampling). AI and robotics map naturally onto these stages, and we formalise this as a two-part stabilisation framework in Section 2.3 (interpretive vs confirmatory stabilisation) [31,33–37,43]. These stabilisation mechanisms differ, and so do their governance needs. AI requires continuous monitoring for drift, bias, and calibration [10,11,20]; robotics requires procedural auditability, operator training control, calibration assurance, and reliable integration of navigation with imaging feedback [33–35,37,43].

Empirical studies and emerging consensus suggest that AI is most valuable when embedded within workflow rather than used as a stand-alone "algorithm". Professional consensus and survey evidence on radiologists' expectations position AI primarily as an augmentative technology that supports clinical interpretation rather than replacing clinician judgement [3,11–13]. In radiology, successful deployment depends on interoperability with PACS/RIS, role clarity (e.g., triage vs second reader), and mechanisms to manage medico-legal responsibility, particularly in population screening contexts where small effects can compound at scale [14,39–41]. Real-world implementation evidence further suggests that AI-supported screening workflows can improve programme-level metrics when deployed under routine delivery at scale [38]. Similarly, robotics adoption depends on procedural volume, learning curves, and institutional readiness to standardise protocols, because the measurable value of robotics is often expressed through reproducibility proxies (e.g., fewer needle adjustments; improved targeting accuracy) rather than population-level outcomes [33–35].

This manuscript therefore moves beyond algorithmic performance to examine diagnostic infrastructure: how AI and robotics reshape end-to-end diagnostic pathways, and what system-level requirements enable safe scaling. We synthesise peer-reviewed evidence (2018–2026) and combine it with public regulatory and market landscape signals. Europe is examined because of population-scale screening programmes, mature regulatory frameworks under the EU Medical Device Regulation (MDR), and extensive evidence on imaging AI deployment [14,47]. Saudi Arabia is analysed as an anchor case for the Gulf region due to documented digital health infrastructure, emerging imaging AI integration signals, and regulator-led governance initiatives [26,29,30]. Finally, we include an illustrative industry-academic case (Breast-SlimView, Hera-Mi) to demonstrate how a workflow-embedded interpretation-support product can be described under MDR using publicly available regulatory and product documentation [49]. Together, these sources support an infrastructure-oriented account of adoption: what stabilises diagnostic practice, what remains fragile, and what governance controls are essential.

1.1. Contributions of this study

This study makes four original contributions. First, it introduces a diagnostic framework distinguishing interpretive stabilisation (primarily enabled by AI) from confirmatory stabilisation (primarily enabled by robotics). Second, it constructs a reproducible public-evidence map of peer-reviewed research (2018–2026) linking technical capability to workflow and governance features across AI interpretation and robotic confirmation. Third, it integrates a regulatory landscape signal for European radiology AI availability based on a public analysis reporting the number of CE-marked radiology AI products as of October 2024 [48]. Fourth, it presents a comparative analysis of diagnostic infrastructure adoption signals using Europe and Saudi Arabia as case examples grounded in public regulatory and institutional documentation [26,29,30,47,48]. The conflict-managed illustrative example (Breast-SlimView) is included only where supported by peer-reviewed and publicly available documentation [49,50].

2. Methods: Reproducible public-evidence integrative analysis

We conducted a structured integrative narrative review combined with a reproducible public-evidence analysis to examine how AI-supported interpretation and robotics-supported confirmation reshape end-to-end cancer diagnostic pathways. This is an integrative narrative review supported by structured database searches; it is not a registered systematic review, and no quantitative meta-analysis was performed. The review covers literature published from January 2018 to March 2026 (capture window). Database searches were executed between October 2025 and March 2026 and updated prior to finalisation (execution window). The primary focus was on peer-reviewed studies published between January 2018 and March 2026; however, a small number of earlier foundational peer-reviewed references were retained selectively where they provided essential conceptual, technical, or translational context.

For clarity, regulatory guidance, institutional reports, and vendor documentation are cited only as contextual "signals" and were not included in the PRISMA-counted peer-reviewed evidence stream.

2.1. Search strategy and information sources

We searched major academic and scientific bibliographic databases for studies on (i) artificial intelligence in cancer imaging, including detection, triage, risk stratification, and workflow support, and (ii) robotics and image-guided diagnostic confirmation, including biopsy targeting, sampling reproducibility, and procedural variance. Core searches were conducted in MEDLINE (via PubMed), Scopus, Web of Science Core Collection, and IEEE Xplore. Search strings combined keywords and, where available, controlled vocabulary. Core concepts covered cancer and oncology, imaging and radiology modalities, artificial intelligence methods, and robotics or image-guided biopsy/navigation.

To broaden coverage across clinical and engineering literature, these searches were supplemented by targeted hand-searching of leading journals represented in the reference set (e.g., Radiology, European Radiology, The Lancet Digital Health, Nature Medicine, Cancer Imaging, Cardiovascular and Interventional Radiology, The British Journal of Radiology, Diagnostics, Respiration, and Chest), together with Google Scholar and arXiv searching. We also performed forward and backward citation chasing and screened reference lists of high-relevance reviews, consensus statements, and regulatory or policy documents, including public regulatory guidance and institutional documentation used in the case-study sections.

Full database search strings for MEDLINE/PubMed, Scopus, Web of Science Core Collection, and IEEE Xplore, together with supplementary source search procedures (Google Scholar, arXiv, and citation chasing), are provided in Appendix B. We focused on peer-reviewed studies published between January 2018 and March 2026, with database searches conducted up to March 2026, to capture contemporary AI and robotics systems and deployment contexts.

In this manuscript, "adoption/infrastructure signals" refers to publicly observable indicators of deployment readiness, such as integration pathways, governance structures, service activation, regulatory positioning, and operational workflow arrangements. These signals support contextual

interpretation but are not treated as evidence of clinical effectiveness unless corroborated by peer-reviewed outcome evidence.

2.2. Eligibility criteria and study selection

We included peer-reviewed studies in oncology imaging or image-guided diagnostic confirmation that reported at least one of: (a) clinical diagnostic performance with sufficient methodological detail; (b) workflow or operational endpoints (e.g., reading time, recall stability, triage time, inter-reader agreement); or (c) procedure-linked confirmation endpoints (e.g., targeting error, needle adjustments, sampling adequacy, reproducibility). We excluded purely technical papers without clinical pathway relevance, studies without reproducible methods, and non-oncology clinical studies. A small number of non-oncology peer-reviewed exemplars were retained selectively where they illustrated transferable workflow governance patterns relevant to imaging AI deployment.

Two-stage screening (title/abstract then full text) was applied using prespecified criteria. Because this work is infrastructure-oriented and multi-domain, we did not restrict by study design; instead, study design and evidentiary strength were captured explicitly in the evidence map and used to weight interpretation. Screening and evidence-map extraction were performed by K. Chebib; M. Tardy independently reviewed a purposive subset of extracted items to check internal consistency of domain interpretation. Formal inter-rater agreement statistics were not a design objective for this infrastructure-oriented synthesis.

Study identification and selection are summarised in a PRISMA-style flow diagram (Fig. 1). The flow reports the peer-reviewed literature stream only; public regulatory and institutional sources were mapped separately and are not included in the selection counts.

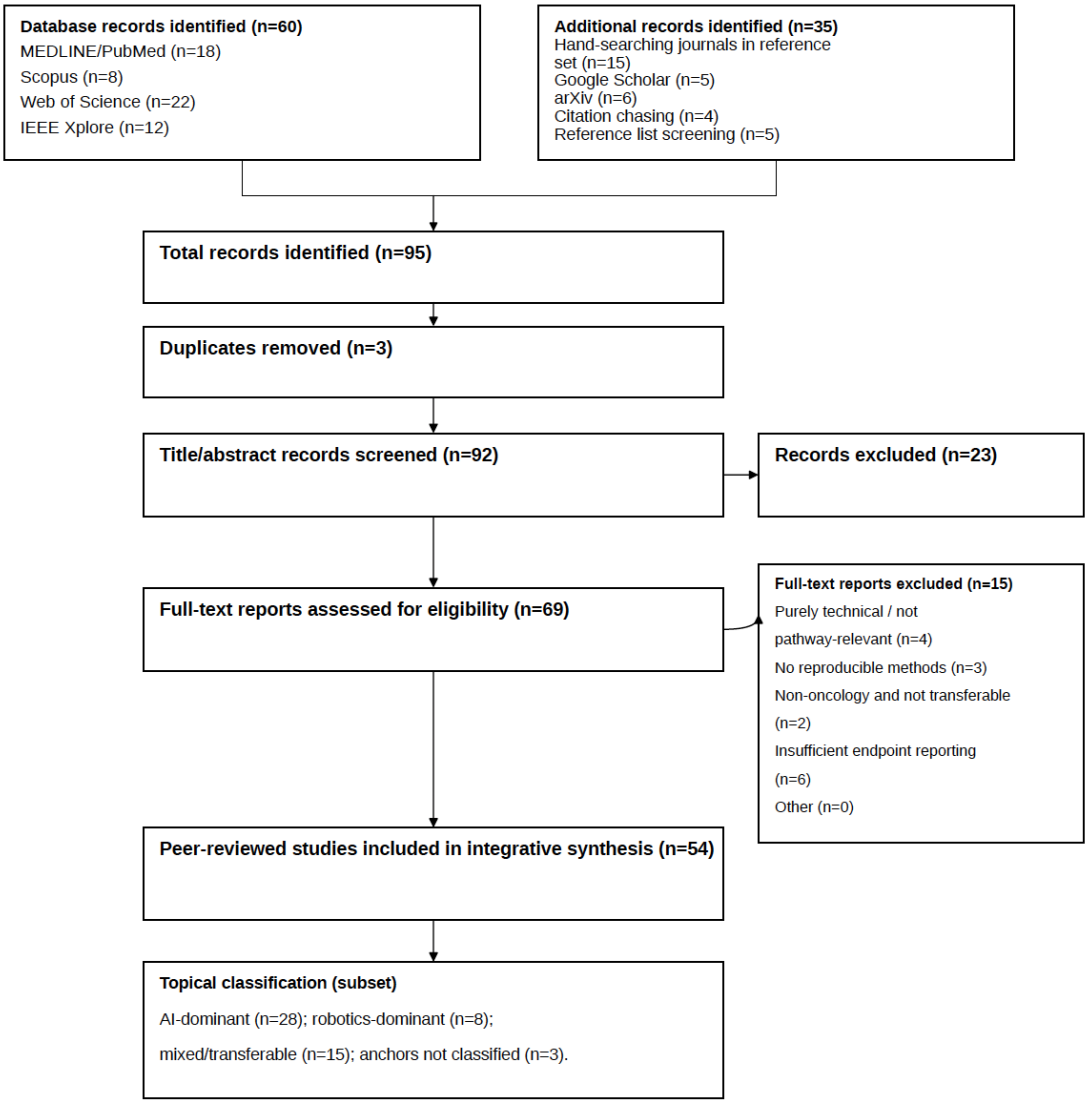


Fig. 1. PRISMA-style flow of peer-reviewed evidence identification and study selection. Records identified (n=95) → duplicates removed (n=3) → screened (n=92) → excluded at title/abstract (n=23) → full texts assessed (n=69) → full texts excluded (n=15) → included (n=54). The flow reports the peer-reviewed literature stream only; public regulatory and institutional sources were mapped separately and are not included in the selection counts.

2.3. Data extraction and evidence mapping

We extracted data into a structured evidence map capturing: (i) cancer type and clinical context, (ii) imaging modality and diagnostic task, (iii) AI model or robotics system characteristics, (iv) workflow role (triage, second reader, targeting assistance), (v) reported outcomes (accuracy, efficiency, reproducibility proxies), and (vi) governance and adoption attributes (integration with PACS/RIS, accountability model, monitoring approach, regulatory framing).

We distinguish two complementary stabilisation mechanisms across the diagnostic pathway. Interpretive stabilisation refers to AI-supported reduction of variance in image interpretation (e.g., detection, prioritisation/triage, quantification, and reader consistency). Confirmatory stabilisation refers to robotics- and image-guidance-supported reduction of variance in diagnostic confirmation (e.g., targeting, access, sampling reproducibility, and procedural execution), linking imaging suspicion to verifiable tissue acquisition or staging endpoints. We apply this framework throughout the manuscript to interpret workflow and governance requirements alongside clinical evidence.

2.4. Quality and risk-of-bias appraisal

To avoid overstating conclusions from heterogeneous evidence, we used domain-appropriate appraisal lenses: QUADAS-2 for diagnostic accuracy studies, PROBAST for prediction model studies, and ROBINS-I or Cochrane RoB 2 domains where applicable for comparative clinical studies. For AI reporting quality and transparency, we assessed alignment with CLAIM and, where relevant, CONSORT-AI/SPIRIT-AI elements. Appraisal was used to weight the narrative synthesis (e.g., emphasising externally validated, prospectively evaluated, or routinely deployed evidence) rather than to generate a single numerical score.

2.5. Public-evidence triangulation and case study selection

To contextualise adoption beyond the literature, we triangulated peer-reviewed findings with public regulatory and market signals, including EU MDR regulatory framing and a European landscape analysis of CE-marked radiology AI products [48]. Europe and Saudi Arabia were selected as analytically grounded cases representing different governance and scaling conditions. Case-study sources were restricted to publicly available materials (regulator guidance, institutional reports, official announcements, and peer-reviewed publications) [26,29,30,47,48].

An illustrative industry-academic case involving Breast-SlimView (Hera-Mi) is included solely to demonstrate how a workflow-embedded interpretation-support product can be described within this infrastructure framework. To mitigate conflicts of interest, we limit factual claims in the case to publicly available regulatory and product documentation and do not use proprietary performance data or unpublished operational dashboards as evidence.

Box 1. Rules for public-evidence triangulation and COI-managed case inclusion

To avoid conflating deployment publicity with clinical evidence, public sources were used under the following constraints:

- Adoption signals only: vendor or institutional announcements, procurement notes, and product pages were treated as signals of uptake (what was deployed), not as evidence of effect.
- Evidence anchoring: claims about diagnostic accuracy, workload, or patient impact were supported by peer-reviewed studies or formal evaluations; public sources were used only to describe governance context and implementation configuration.
- Traceability: every quantitative claim includes the underlying denominator (e.g., number of studies, examinations, reads, or procedures) and a direct citation to the primary source.

• COI-managed inclusion: where an illustrative case involved author affiliation, claims were restricted to publicly available documentation; no proprietary data were used, and interpretations were explicitly framed as analytical exemplars rather than effectiveness evaluations.

2.6. Ethics and reporting

This study used publicly available sources and did not involve access to identifiable patient data; formal ethics review was therefore not required. Given the objective of an infrastructure-oriented narrative synthesis spanning multiple domains, we prioritised transparent and reproducible reporting over systematic-review protocol registration. To support transparency, we provide eligibility criteria, a structured evidence map, and a PRISMA-style flow diagram summarising peer-reviewed study identification and selection (Fig. 1). Key quantitative claims are linked to specific sources.

3. Technical Foundations: Complementary Diagnostic Levers

3.1. AI as Interpretive Stabilisation

Deep learning systems can improve cancer diagnosis by enhancing lesion detection, malignancy risk estimation, and triage across imaging modalities [1,2,16–19]. Their primary contribution is reducing cognitive variability and supporting consistency in high-volume or low-prevalence contexts. Evidence from thoracic imaging illustrates these capabilities operationally, including lung nodule detection and lung cancer risk assessment in low-dose CT screening contexts [16,17]. AI can also influence upstream image quality and interpretability through reconstruction and enhancement approaches, including deep-learning–based methods in CT and MRI, and through dose-reduction and reconstruction pipelines that affect task-based imaging performance [4].

Although not oncology-specific, clinically deployed retinal AI referral systems provide an early, widely cited example of how imaging AI can be integrated into real-world diagnostic workflows with explicit referral thresholds and escalation pathways; we include this as an infrastructure archetype rather than cancer evidence [37].

3.2. Robotics as Confirmatory Stabilisation

Robotics contributes at the confirmatory stage of the diagnostic pathway. When imaging identifies a suspicious lesion, diagnostic certainty often depends on tissue sampling or minimally invasive confirmation procedures. These procedures introduce another source of variability: procedural variance. Operator-dependent differences in needle placement accuracy, access planning, lesion targeting, and sampling adequacy can affect diagnostic outcomes and may lead to repeat procedures or false-negative confirmation [33–35].

Robotic assistance in image-guided biopsy and navigational procedures addresses this procedural variance by improving targeting precision, stabilising movement, and enabling reproducible access paths, particularly for anatomically challenging lesions [33,34]. Image-guided intervention and navigation technologies provide the technical substrate for robotics-assisted confirmation, including image-to-device registration, targeting and trajectory planning, and reproducible

instrument guidance under imaging constraints [31,33,36,43]. Contemporary oncology-focused reviews describe how image-guided robotics can standardise components of interventional workflows and reduce the need for repeated needle manipulation in CT-guided and related pathways [31,33,34]. In this sense, robotics provides confirmatory stabilisation: reducing variability in the physical steps that convert imaging suspicion into diagnostic confirmation [33–35].

These complementary mechanisms of variance reduction across interpretive and confirmatory stages are illustrated in Figure 2.

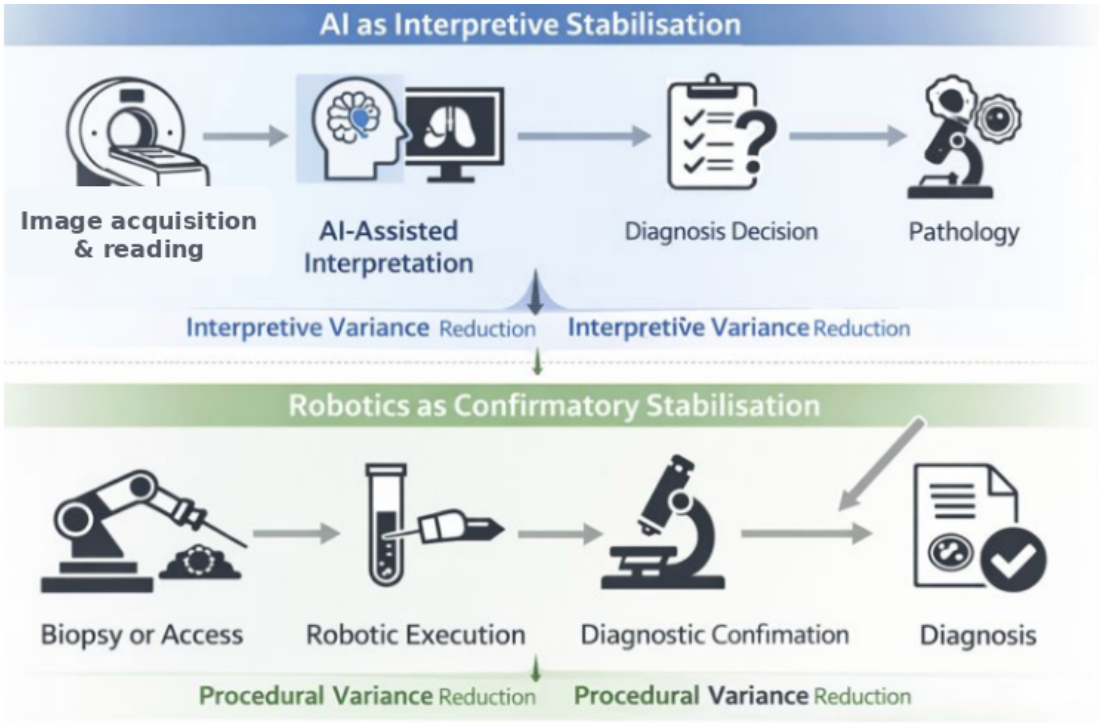


Figure 2. Variance reduction mechanisms of AI and robotics in cancer diagnosis.

AI functions as **interpretive stabilisation** by reducing inter- and intra-reader variability during image interpretation (e.g., detection support, prioritisation/triage, consistency under workload). Robotics functions as **confirmatory stabilisation** by reducing procedural variability during image-guided diagnostic confirmation (e.g., stabilised targeting, reduced needle adjustments, improved reproducibility of sampling). Diagnostic value is maximised when both variance-reduction mechanisms operate across the end-to-end diagnostic pathway.

4. Clinical Performance in Real-World Settings

4.1. AI performance degradation outside controlled environments

While many AI models demonstrate strong performance in retrospective or controlled evaluations, real-world diagnostic environments are more heterogeneous. Differences in scanner vendors, protocols, patient populations, prevalence, and workflow context can reduce model generalisability and lead to performance drift over time [10,11,20]. In cancer imaging, this can manifest as decreased sensitivity in underrepresented subgroups, false positives in atypical presentations, or reduced specificity in low-prevalence screening settings [16].

This discrepancy between development performance and deployment performance highlights the importance of governance: continuous monitoring, calibration assessment, and human override mechanisms to ensure safe and sustained diagnostic benefit [10,11,20,23–25].

4.2. Procedural variance as a diagnostic bottleneck

Robotics-supported diagnostic confirmation exhibits a different performance profile. Procedural outcomes are influenced by operator expertise, anatomy, lesion accessibility, and equipment calibration. Robotics can reduce targeting error and limit repositioning in image-guided diagnostic procedures, including CT-guided percutaneous needle placement [33] and navigational bronchoscopy pathways for peripheral pulmonary lesions [34]. However, robotics does not eliminate the need for clinical judgement, nor does it fully standardise procedures across institutions without training and quality assurance mechanisms.

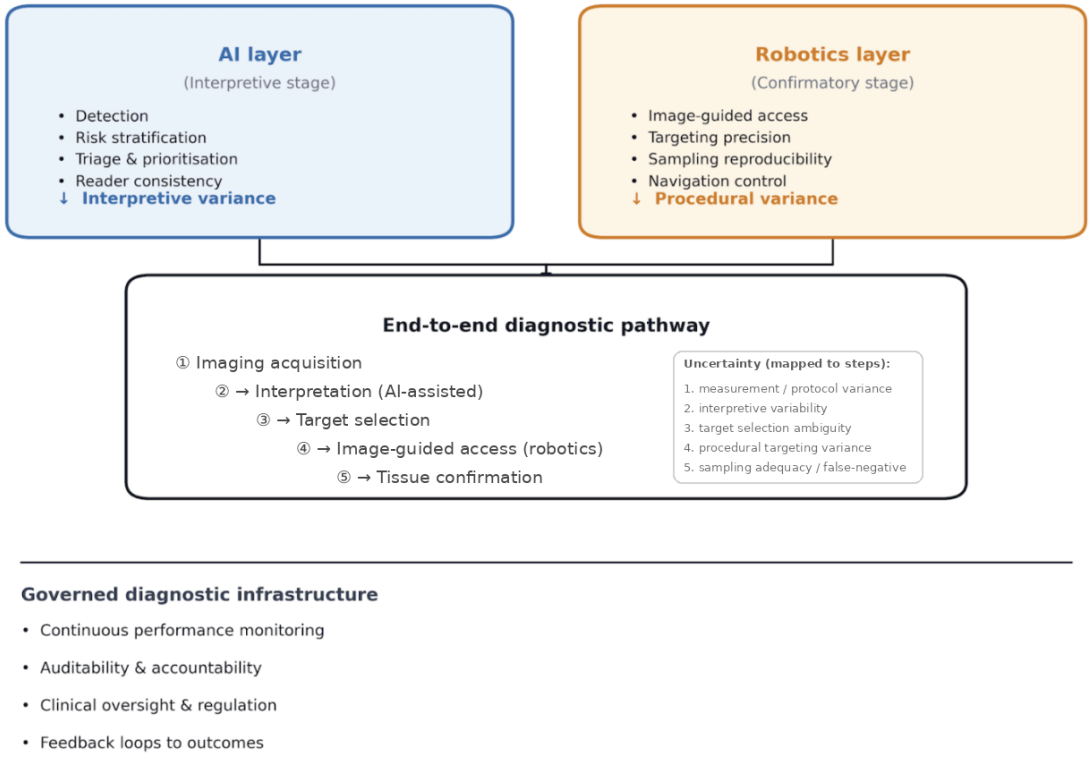
Robotics evaluation is therefore often anchored to procedural precision and reproducibility metrics rather than large-scale population-level outcomes. This reflects robotics' clinical value at the individual procedure level, where improvements in targeting reliability can reduce repeat procedures, complications, and diagnostic delay in cases requiring tissue confirmation [33–35].

5. Diagnostic Integration and Adoption

This section integrates findings from the integrative literature review and the structured public-evidence map constructed for this study (Tables 1–3), enabling a workflow-level and system-level analysis of how artificial intelligence and robotics are adopted in cancer diagnosis. Rather than evaluating technologies in isolation, the analysis focuses on where diagnostic uncertainty arises, how it propagates across the diagnostic pathway, and which technological interventions can reduce that uncertainty in a durable and governable manner.

Figure 2 describes the mechanistic sources of diagnostic variance and the corresponding technological levers, whereas Figure 3 maps uncertainty points across an end-to-end diagnostic workflow and shows where stabilisation mechanisms are applied operationally. Together, these figures distinguish the type of uncertainty (Figure 3, markers 1–5) from the mechanisms through which it is mitigated in practice (Figure 2).

Figure 3. End-to-end diagnostic pathway and complementary stabilisation roles of AI and robotics



Schematic representation of an end-to-end diagnostic pathway illustrating how artificial intelligence reduces variance across image acquisition support, triage, and reporting, while robotics primarily reduces procedural variability during image-guided diagnostic confirmation (e.g., biopsy).

Note: Conceptual schematic; specific workflow steps and outcomes vary by modality, cancer type, and local implementation. Numbered markers indicate uncertainty points along the pathway (measurement/protocol variance, interpretive variability, target selection ambiguity, procedural targeting variance, and sampling adequacy/false-negative confirmation).

5.1. AI integration into diagnostic workflows

The clinical impact of artificial intelligence in cancer imaging is determined less by standalone algorithmic accuracy than by how AI is embedded into routine diagnostic workflows [2,16,17]. Across modalities and clinical contexts, our analysis shows that AI adoption is highest when systems are PACS-native, operate within existing reporting environments, and reduce cognitive burden rather than introduce parallel decision layers [16,17]. Large-scale European screening programmes illustrate this dynamic particularly clearly [38–41]. National breast screening systems process very high volumes of mammographic examinations, creating sustained interpretive workload, visual fatigue, and risk of intra- and inter-reader variability [38–41]. In this context, AI tools deployed as second readers or decision-support layers primarily improve interpretive consistency and throughput, rather than dramatically shifting headline sensitivity or specificity metrics [38–41].

Double reading has demonstrable value in diagnostic radiology by improving reliability and reducing missed findings, providing an operational analogue for how AI may function as a

consistent second-reader layer in screening workflows [14,38–41]. Reported impacts of AI-supported reading include reductions in average reading time per case and stabilisation of recall behaviour, effects that compound at population scale even when per-case gains appear modest [39–41]. Real-world implementation evidence now indicates that AI-supported screening workflows can improve programme-level metrics when deployed within routine delivery at scale [38]. Conversely, AI tools that require clinicians to exit PACS environments, reconcile discordant outputs, or interpret opaque probability scores show markedly lower uptake [12,16,17]. Empirical workflow studies demonstrate that such designs can paradoxically increase reporting time or reduce clinician trust, despite strong standalone algorithmic performance [12,16,17].

From an adoption perspective, AI integration therefore functions as a sociotechnical optimization problem rather than a purely technical one [16,17,25]. Tools succeed when they align with reporting habits, accountability structures, and existing quality assurance processes [25,38–41]. This helps explain why AI deployment has progressed fastest in screening and triage contexts, where interpretive variance is a dominant diagnostic bottleneck, and more slowly in complex, low-volume diagnostic tasks where human contextual reasoning remains central [14,18,19,38–41].

5.2. Robotics integration into diagnostic procedures

Robotics adoption in cancer diagnosis follows a fundamentally different logic. Here, the principal constraint is not interpretive workload but procedural variability in image-guided diagnostic confirmation steps, particularly biopsy. Robotics becomes diagnostically valuable when it improves targeting precision, reproducibility, and sampling reliability without adding friction to procedural workflows [31,33–37]. Evidence from interventional imaging demonstrates that robotic assistance integrated with CT, CBCT, MRI, and ultrasound improves needle placement accuracy and reduces the number of repositioning attempts compared with freehand techniques [33–35]. These effects are especially pronounced in anatomically challenging lesions or small targets, where minor deviations can lead to non-diagnostic samples, repeat procedures, or delayed diagnosis.

European evaluations of robot-assisted CT-guided needle positioning specifically report fewer needle adjustments and improved positioning accuracy versus conventional approaches [33,36]. Implementation reports describe the workflow and governance arrangements required to integrate CT-guided robotics into interventional oncology pathways [43]. Crucially, robotics adoption is highly sensitive to setup complexity, calibration time, and operator training. Systems that require extensive pre-procedure preparation or disrupt established imaging-procedure coordination face resistance, even when accuracy gains are demonstrated. Reviews consistently emphasise that diagnostic value arises when robotics reduces procedural variance without extending procedure time or increasing cognitive overhead for operators [31,33,35,37].

In contrast to AI, robotics adoption remains centre-based and selective, reflecting capital cost, training requirements, and minimum procedural volume thresholds. However, where deployed, robotics addresses a diagnostic bottleneck that AI alone cannot resolve: the reliable translation of imaging suspicion into tissue confirmation. This distinction underpins the complementary rather than substitutive relationship between AI and robotics in cancer diagnosis.

5.3. AI and robotics as diagnostic infrastructure

At scale, artificial intelligence and robotics should be understood not as discrete tools but as components of diagnostic infrastructure. Both require longitudinal performance monitoring, auditability, and governance comparable to other safety-critical diagnostic technologies.

For AI, this includes ongoing surveillance for dataset shift, drift in performance across subpopulations, and changes in clinical practice [10,11,20]. For robotics, governance focuses on procedural safety, calibration integrity, operator competency, and outcome tracking. In both cases, diagnostic value depends on institutional capacity to monitor performance over time, respond to deviations, and integrate feedback into clinical governance frameworks.

Health systems that treat AI and robotics as infrastructure, embedding them into quality assurance, audit, and reporting processes, report more sustainable adoption than those treating them as pilot technologies. This framing is particularly relevant for cancer diagnosis, where false negatives, repeat procedures, or inconsistent confirmation pathways propagate downstream into treatment decisions, patient outcomes, and system-level trust.

The workflow-level and adoption patterns described in this section are synthesised quantitatively and comparatively in Tables 1–3.

Table 1 characterises how diagnostic performance gains differ by diagnostic stage; Table 2 captures workflow and outcome effects; and Table 3 contrasts adoption maturity and value drivers across AI and robotics, reinforcing the conclusion that durable diagnostic improvement arises from coordinated variance reduction across the diagnostic pathway, rather than isolated technological optimisation.

Table 1. Diagnostic performance metrics across AI interpretation and robotic confirmation

Domain	Stage	Metric	Reported effect	Evidence	Sources
AI – cancer imaging (general)	Interpretation	Sensitivity / specificity / AUC	Strong performance in controlled evaluations; clinical value depends on workflow embedding.	Systematic review with meta-analysis	[1,2,16]
AI – triage (e.g., urgent imaging pathways)	Timeliness	Time to specialist review / transfer time	Reduced.	Observational implementation studies (retrospective/prospective service evaluations)	[18,19]

Domain	Stage	Metric	Reported effect	Evidence	Sources
Robotics – image-guided biopsy (CT/US/MRI-assisted)	Confirmation	Needle placement accuracy	Improved; fewer adjustments.	Comparative procedural studies (typically single-centre) plus systematic or narrative reviews	[31,33–37]
Robotics – biopsy pathways	Confirmation	Repeat / non-diagnostic sampling proxies	Reduced variability; fewer re-positionings and improved targeting consistency.	Comparative procedural studies plus reviews (workflow/procedur level endpoints)	[31,33–37]

Interpretation: Table 1 synthesises indicative performance and reliability metrics across AI-supported interpretation and robotics-supported diagnostic confirmation. AI primarily contributes at the interpretive stage (consistency/timeliness), whereas robotics contributes at the confirmatory stage (targeting precision and sampling reproducibility).

Abbreviations: AUC, area under the receiver operating characteristic curve; CT, computed tomography; US, ultrasound; MRI, magnetic resonance imaging.

Note: Evidence types are heterogeneous across domains. AI interpretation studies more often report diagnostic or programme-level performance outcomes, whereas robotics evidence is more commonly based on procedural, single-centre, or feasibility-oriented studies using targeting, positioning, or workflow endpoints. Accordingly, comparisons in this table are intended to illustrate stage-specific patterns of diagnostic value rather than direct equivalence in evidence maturity.

Table 2. Workflow and outcome impacts across AI and robotics in cancer diagnosis

Technology	Context	Workflow metric	Observed impact	Diagnostic relevance	Sources
AI	Screening / high-volume reporting	Reading workflow integration and reporting efficiency proxies	Improved workflow efficiency when PACS-aligned; supports second-reader workflows.	Throughput and consistency	[12,14,16,17,38–41]
AI	Acute neuroimaging networks (deployment model)	Time to specialist review / transfer time	Reduced.	Earlier diagnosis or routing	[19]

Technology	Context	Workflow metric	Observed impact	Diagnostic relevance	Sources
Robotics	Image-guided biopsy	Needle adjustments / targeting reproducibility	Reduced adjustments; increased reproducibility.	Sampling reliability	[31,33–37]
AI + robotics	End-to-end pathway (conceptual synthesis)	Repeat procedures / delay proxies	Reduced when variance is reduced at both interpretation and confirmation stages.	End-to-end reliability	[16,17, 33, 34,43]

Interpretation: Table 2 highlights that the practical value of AI and robotics lies in how they reconfigure diagnostic workflows, not in stand-alone model or device performance metrics. AI can create system-level efficiency gains when embedded into reading pathways (e.g., triage or AI-supported double reading), while robotics can stabilise the confirmation stage by standardising access and sampling. In organised breast screening, the MASAI trial reported a 44.3% reduction in screen-reading workload (46,345 vs 83,231 reads) under an AI-supported configuration with safety monitoring, alongside non-inferior screening accuracy; complementary trial and real-world reports describe how performance and operational effects depend on programme rules and deployment context [38–40]. In image-guided biopsy, robotic assistance studies typically quantify stabilisation via fewer needle re-positionings/trajectory corrections and improved targeting precision rather than downstream patient outcomes [31,33–36].

Table 3. Comparative synthesis of adoption maturity and diagnostic roles for AI versus robotics

Dimension	AI in imaging	Robotics in diagnosis
Primary diagnostic role	Interpretation	Confirmation
Maturity	Established in selected imaging use cases [15,38–41]	Selective centres (interventional radiology and biopsy pathways) [33–35,43]
Evidence scale	Often assessed in large cohorts and programme studies [38–41]	Primarily assessed in procedure-level studies and series [33–35,43]
Key adoption barrier	Workflow integration and trust; performance drift and subgroup bias; lifecycle governance requirements including monitoring, change control, auditability, and accountability [10,11,20,25,42,47,52,66–68]	Cost, training, integration, QA and maintenance [33–35,43]

Dimension	AI in imaging	Robotics in diagnosis
Diagnostic value driver	Reading variance reduction and routing or triage efficiency [14,18,19,38–41]	Procedural variance reduction (targeting precision, reproducibility) [33–35,43]

Interpretation: Table 3 summarises the contrasting adoption dynamics of AI and robotics in cancer diagnosis. AI is established in selected imaging use cases and is often evaluated at programme scale, including large cohorts and screening studies [15,38–41]. This relative maturity is also reflected in public regulatory and market-visibility signals. In the United States, the FDA maintains a public list of Artificial Intelligence-Enabled Medical Devices, which functions as a transparency signal for authorised AI-enabled devices across specialties rather than a direct measure of routine clinical uptake [54]. A 2025 comparative analysis of radiology AI similarly reported a substantially larger publicly visible authorised market in the United States than in Europe, while noting that registry or approval counts should not be interpreted as equivalent to validated clinical adoption [63]. In China, a 2025 review likewise described increasing NMPA approvals over time, with radiology representing a large share of approved AI-enabled devices [65]. By contrast, robotics adoption remains more selective and centre-based, concentrated in interventional radiology and biopsy pathways, and is predominantly supported by procedure-level studies and clinical series rather than population-scale programmes [33–35,43]. The adoption constraints also differ. For AI, the main constraints relate to workflow integration, human-AI interaction, subgroup robustness, performance drift, and lifecycle governance [10,11,20,23–25,42,47,52,66–68]. For robotics, the main constraints relate to capital cost, operator training, integration into procedural workflows, and ongoing quality assurance and maintenance [33–35,43]. Critically, the diagnostic value drivers differ as well: AI primarily reduces reading variance and can improve routing or triage efficiency [14,18,19,38–41], whereas robotics primarily reduces procedural variance through improved targeting precision and reproducibility [33–35,43]. Together, these differences help explain why neither technology alone is sufficient to optimise end-to-end cancer diagnosis and why durable benefit depends on workflow design, governance, and institutional learning [25,33–35,43,62].

6. Ethical, Legal, and Governance Considerations

The clinical value of AI and robotics in cancer diagnosis depends not only on technical performance but also on the ethical, legal, and governance frameworks within which these technologies operate [25,47,51,52,66–68]. Unlike many decision-support tools, AI- and robotics-enabled diagnostic systems intervene at critical points in the diagnostic pathway, interpretation and confirmation, where errors can propagate downstream into treatment decisions, patient outcomes, and system-level trust [12,23,33–35,42]. As such, governance is not an adjunct to innovation but a precondition for diagnostic safety, legitimacy, and sustainability [25,47,51,52,66–68].

6.1. Ethical and legal risks in AI-supported diagnostic interpretation

AI systems deployed in cancer imaging introduce ethical risks related to bias, false reassurance, and performance drift [12,23,42,67]. Bias arises when training and validation datasets under-represent demographic groups, scanner vendors, acquisition protocols, or disease phenotypes, leading to systematic underperformance in specific populations [12,23,66]. In screening and other high-volume settings, such biases can remain clinically invisible at the individual-case level while generating population-level inequities [23,42,66].

False reassurance constitutes a second ethical risk. When AI outputs (e.g., heatmaps, probability scores, triage flags) are embedded into worklists or prioritisation systems, they can be perceived as authoritative and may shift clinician attention and confidence in subtle ways, especially in ambiguous cases [23,42]. The absence of clear uncertainty communication, defined escalation rules, and local governance on appropriate reliance exacerbates this risk [25,64,66,67,68].

Performance drift further complicates ethical and legal accountability. Changes in scanner hardware, imaging protocols, disease prevalence, referral thresholds, or case-mix can degrade model performance over time, rendering one-off validation insufficient [10,11,20]. From a legal and regulatory standpoint, this motivates lifecycle obligations, controlled updates, change control planning, post-deployment monitoring, and revalidation, so that software-enabled medical device functions remain safe and effective under real-world conditions [47,52,68]. Governance frameworks must therefore extend beyond pre-market validation to include continuous monitoring, revalidation, and auditability [25,47,52,68].

6.2. Accountability and professional responsibility in AI-assisted diagnosis

Despite increasing automation, clinicians retain responsibility for diagnostic decisions. However, AI systems increasingly influence attention allocation, prioritisation, and diagnostic confidence, blurring practical boundaries between human judgement and algorithmic suggestion [42,66]. This raises questions about accountability, documentation, and medico-legal exposure, particularly when AI is integrated as triage, second reading, quantification, or risk stratification within routine workflows [25,47,51,52,64,66–68].

Accordingly, governance requires explicit role definition. This includes what the AI is authorised to do, what it is not authorised to do, how disagreements between clinician interpretation and AI output are documented and resolved, and how audit trails are maintained for retrospective review, quality assurance, and organisational learning [25,47,52,66–68]. Documentation practices should make AI involvement sufficiently transparent to support audit and incident review while remaining workable in high-throughput diagnostic environments [25,47,52,66–68]. In regulated settings, accountability is strengthened when organisations treat AI as safety-critical diagnostic infrastructure, subject to quality assurance, monitoring, incident response, and controlled updates, rather than as a static "add-on" [25,47,52,68].

6.3. Ethical and legal risks in robotics-supported diagnostic confirmation

Robotics used in image-guided biopsy and diagnostic procedures introduces a distinct risk profile centred on procedural and mechanical variability rather than interpretive error [33–35,37,43]. Ethical concerns include calibration accuracy, mechanical reliability, and the interaction between operator skill and device behaviour. Errors at this stage may result in non-diagnostic samples, repeat procedures, or procedural complications, directly affecting patient safety and diagnostic confidence [33–35].

Accountability in robotics-assisted diagnosis is particularly complex because robotic systems physically mediate access to tissue. Current practice generally assigns responsibility to the operating clinician. However, increasing levels of automation and image fusion raise questions about shared accountability across clinicians, institutions, and device manufacturers [35,37]. Governance frameworks must therefore address not only device approval but also operator training, credentialing, learning curves, and institutional oversight structures [35,37,58–62].

Evidence from interventional robotics reviews underscores that diagnostic gains depend on consistent calibration, procedural standardisation, and outcome tracking [31,33–35,43]. Robotics delivers value when embedded within institutional quality assurance systems that treat it as diagnostic infrastructure, subject to the same audit and performance review processes as other safety-critical diagnostic technologies [34,43,62].

6.4. Emerging ethical and governance challenges in micro-/nanorobotics

Micro- and nanorobotics introduces additional ethical and governance challenges that extend beyond those associated with current diagnostic technologies. Proposed applications, such as in situ sensing, tumour microenvironment mapping, or micro-scale interaction with cancer tissue, raise unresolved questions regarding biocompatibility, clearance pathways, controllability, and long-term biological effects [32]. Unlike macro-robotics, which operate under direct clinician control, micro-/nanorobotic systems may function semi-autonomously within the body, complicating oversight and intervention [32]. These characteristics are likely to challenge existing approaches to device evaluation, risk assessment, and clinical oversight, particularly where diagnostic agents are distributed, mobile, or degradable [32]. They also imply more demanding consent and governance requirements than those associated with conventional imaging or procedural tools [32]. Given these uncertainties, micro-/nanorobotics should presently be regarded as an experimental diagnostic frontier, confined to staged research and validation pathways supported by robust preclinical evidence and stringent ethical oversight [32]. Premature clinical deployment would risk outpacing the governance structures needed for safe and trustworthy use [32].

The organisational implications of these governance requirements are considered further in Section 8.2, where accountability is examined at the level of clinical workflow, role allocation, and institutional oversight.

6.5. Regulatory alignment across major jurisdictions

As AI and robotics become embedded within cancer diagnostic pathways, governance considerations increasingly intersect with formal regulatory regimes. While regulatory frameworks differ across jurisdictions, a clear convergence is evident toward risk-based oversight models that emphasise post-market monitoring, human accountability, and system-level safety rather than static pre-market validation alone [47,51,52,60,64].

In Europe, AI and robotics used in diagnostic imaging fall primarily under the EU Medical Device Regulation (MDR), with additional relevance from the EU Artificial Intelligence Act (AI Act) [47,51]. MDR requires clinical evidence, conformity assessment, and post-market surveillance, explicitly recognising the need for ongoing vigilance [47]. For AI, this aligns with concerns around bias and performance drift; for robotics, MDR focuses on device safety and intended use while situating procedural execution within professional accountability frameworks [25,34,47]. This reinforces the treatment of robotics as diagnostic infrastructure requiring institutional quality assurance [34,47].

In the United States, the Food and Drug Administration (FDA) has adopted a lifecycle-oriented approach to software as a medical device (SaMD), acknowledging that adaptive and learning systems cannot be governed solely through pre-market approval. Emphasis is placed on real-world performance data, change management, and transparency around updates [52,61]. Robotics used in diagnostic confirmation is similarly regulated through device pathways, with accountability anchored in clinician oversight and institutional credentialing [35,37,47,60].

Emerging Gulf regulatory approaches, particularly in Saudi Arabia, reflect a comparable risk-based logic within more centralised health governance structures. The Saudi Food and Drug Authority (SFDA) emphasises registration, clinical justification, and post-market oversight for AI-enabled diagnostic tools [30]. Although regulatory detail is less publicly granular, the close coupling between regulator, national health operators, and tertiary centres facilitates system-level governance of AI and robotics as part of national diagnostic infrastructure [30]. For robotics, regulatory oversight is complemented by centre-level readiness requirements, including training, procedural protocols, and audit mechanisms [30,35,37].

Across jurisdictions, three common principles emerge: human accountability remains central, longitudinal performance monitoring is essential, and emerging technologies are subject to staged validation [47,51,52,60,64]. These regulatory patterns reinforce a central conclusion of this study: AI and robotics improve cancer diagnosis only when technical capability, clinical workflow, and governance evolve in parallel [25,34,47,51,52]. Regulation does not merely constrain innovation; it shapes how diagnostic technologies mature into trusted, sustainable infrastructure [47,51,52,60,64].

7. Continuous Learning and Sustainability

The long-term diagnostic value of artificial intelligence and robotics in cancer imaging depends not on peak performance at the point of regulatory approval or initial deployment, but on their capacity to remain reliable, interpretable, and clinically aligned as diagnostic environments evolve. Cancer diagnostic pathways are characterised by continuous change: imaging hardware is upgraded, acquisition protocols drift, screening populations age, referral thresholds shift, and clinical workflows are reorganised in response to policy, workforce pressure, and technology adoption. Static AI models and fixed procedural configurations may not remain well-aligned with real-world oncology practice unless supported by ongoing monitoring, quality assurance, and controlled adaptation [10,11,20,25,34].

7.1. Continuous performance oversight in AI-supported interpretation

For AI systems deployed in diagnostic interpretation, sustainability requires longitudinal performance oversight linked to clinical outcomes, rather than reliance on pre-market validation alone. Even well-validated models can experience clinically meaningful performance degradation when exposed to new scanners, revised protocols, prevalence shifts, or demographic change [10,11,20].

In breast screening, small absolute changes in false-negative or false-positive rates can translate into large programme-level effects. Sustainable deployment therefore depends on continuous monitoring of metrics such as recall stability, interval cancer rates, and subgroup performance across breast density, age, and acquisition modality, coupled to clear escalation and recalibration pathways.

Importantly, continuous learning in AI-supported interpretation is typically institutionally mediated rather than autonomous. In regulated settings, model updates are introduced through controlled versioning with documented performance deltas and revalidation where required, preserving clinician trust and regulatory accountability while still allowing improvement over time.

7.2. Continuous quality assurance in robotics-supported diagnostic confirmation

For robotics used in image-guided diagnostic confirmation, sustainability is driven by procedural quality assurance rather than algorithmic adaptation. Robotic systems do not degrade through dataset shift, but their diagnostic performance is sensitive to calibration accuracy, imaging-robot fusion fidelity, operator experience, and procedural volume.

Comparative studies of CT-guided biopsy robotics report fewer needle repositionings and improved targeting accuracy versus freehand techniques [33,36]. Implementation reports from European centres describe the practical workflow and governance arrangements required to integrate CT-guided robotics into interventional oncology pathways [43]. However, these gains are contingent on consistent calibration and trained operator use. Over time, mechanical wear, software updates, or changes in imaging protocols can subtly affect performance if not actively monitored.

Sustainable robotics deployment therefore requires continuous tracking of procedural proxies for diagnostic reliability, such as non-diagnostic biopsy rates, repeat procedures, targeting deviations, and complication profiles. These metrics are most meaningful when linked to downstream pathological confirmation, closing the loop between imaging suspicion, procedural execution, and diagnostic outcome.

The Saudi Arabia case provides a system-level illustration of this logic. Major tertiary centres such as King Faisal Specialist Hospital & Research Centre report >1,000 robotic procedures annually, and Johns Hopkins Aramco Healthcare reports several hundred procedures per year, volumes used here only as contextual indicators of programme maturity rather than outcome superiority [44–46]. At this scale, sustainability depends on institutional capacity to standardise training, maintain calibration, and integrate robotics into multidisciplinary diagnostic pathways. Robotics is thus treated as part of diagnostic infrastructure, subject to ongoing audit and governance, rather than as a one-off technical upgrade.

7.3. Organisational learning as a prerequisite for sustainability

Beyond individual technologies, sustainable diagnostic improvement depends on the ability of health systems to operate as learning organisations. AI and robotics introduce new forms of variability, new dependencies between human judgement and technical systems, and new failure modes that cannot be addressed through device-level optimisation alone.

Health systems that demonstrate durable adoption, both in European screening programmes and in Gulf tertiary centres, share a common feature: AI and robotics are embedded within existing quality assurance, audit, and governance structures. Performance signals are reviewed in multidisciplinary forums, deviations are investigated, and corrective actions are implemented at the system level [25]. This organisational learning capacity is particularly critical in oncology, where diagnostic errors propagate downstream into treatment selection and patient outcomes.

Conversely, pilot-driven or fragmented deployments that lack institutional feedback loops risk accumulating what may be termed *diagnostic debt*[1]: small, unobserved degradations in performance that compound over time. Continuous learning frameworks mitigate this risk by making diagnostic performance visible, accountable, and correctable.

7.4. Sustainability as a diagnostic property of systems

Taken together, these findings reinforce a central conclusion of this study: sustainability in cancer diagnostic innovation is not a property of AI models or robotic platforms in isolation, but of the systems that govern their use [25,34]. AI delivers durable value when embedded in infrastructures capable of monitoring interpretive behaviour across millions of cases and managing dataset shift and performance drift over time [10,11,20,25]. Robotics delivers durable value when integrated into procedural ecosystems that maintain quality assurance and track targeting accuracy, sampling adequacy, and confirmation outcomes longitudinally [34,43].

[1]We use "diagnostic debt" as an author shorthand by analogy to the "technical debt" metaphor in software engineering; here it denotes accumulated diagnostic backlog/delay and downstream clinical and operational consequences..

Health systems that invest in continuous monitoring, controlled adaptation, and outcome-linked governance are better positioned to convert technological capability into lasting diagnostic improvement [25,34,47,52,60,64]. Those that treat AI and robotics as static, deploy-and-forget solutions risk eroding early gains as clinical environments evolve [10,11,20,25].

In this sense, continuous learning is not an optional enhancement but a defining characteristic of diagnostic infrastructure in the era of AI and robotics [25,52,60,61].

8. Technology Frontiers and Future Directions

8.1. Advanced Robotics and Interventional Imaging

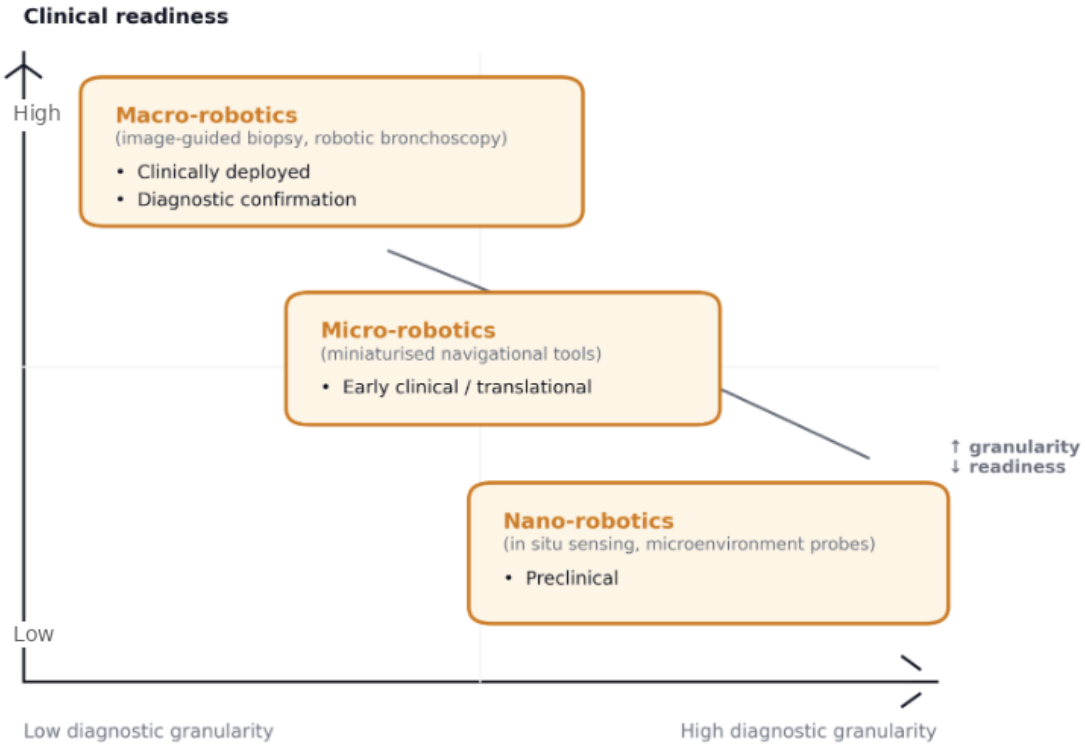
Next-generation robotics systems increasingly integrate AI-assisted planning, real-time imaging feedback, and haptic or autonomous guidance, potentially further reducing operator variance and improving procedural repeatability in image-guided confirmation workflows [35,37].

8.2. Micro-/Nanorobotics as an Emerging Diagnostic Frontier

Micro-/nanorobotics extends robotic assistance from procedure-scale devices to micro- and nano-scale agents. Proposed diagnostic functions include in situ sensing, tumour microenvironment mapping, targeted biomarker detection, and micro-scale interaction with cancer tissue. Most micro-/nanorobotic concepts remain preclinical, with key translational requirements including biocompatibility, clearance pathways, controllability in complex physiological environments, real-time localisation/tracking, and the challenge of evidencing safety and diagnostic benefit under existing medical device frameworks [32].

In this manuscript, micro-/nanorobotics is included to delineate the boundary between near-term diagnostic infrastructure (macro-robotics supporting image-guided confirmation) and longer-horizon diagnostic frontiers. Ethical, legal, and governance implications are discussed in Section 6.4.

Figure 4. Robotics modalities for cancer diagnosis by clinical readiness and diagnostic role.



Interpretation note: Robotic approaches used in cancer diagnosis can be positioned by clinical readiness and the diagnostic granularity they enable. Clinically deployed macro-robotic systems predominantly support image-guided diagnostic confirmation, whereas micro-robotic platforms remain at early clinical or translational stages and nano-robotic concepts are largely preclinical, offering increasingly fine-grained in situ sensing and microenvironment interaction. This framing distinguishes current diagnostic infrastructure from longer-term diagnostic frontiers. Readiness tiers should be interpreted as comparative stages of translational and organisational maturity rather than formal regulatory categories: early readiness denotes proof-of-concept or limited clinical feasibility; intermediate readiness denotes workflow-integrated but still selectively evaluated deployment; and advanced readiness denotes governed, routine use with defined monitoring, quality assurance, and organisational embedding.

9. Case Studies: Balanced Diagnostic Adoption Models

9.1. Europe, Imaging-Based Cancer Diagnosis: System-Level Adoption with Quantitative Signals

Europe offers an instructive model for scaling AI and robotics in cancer diagnosis because adoption is coupled to strong quality assurance norms, risk-based medical device regulation, and organised screening workflows [38–40,47]. In breast screening, prospective and real-world evidence suggests that AI-supported reading can reduce radiologist workload while maintaining screening safety and performance [38–40]. Observed effects are pathway-dependent: they vary with whether AI is used as a triage layer, an independent second reader, or a decision-support overlay within conventional double reading [39–42]. This section therefore treats Europe as a governance-and-workflow setting rather than a single national implementation [38–40,47].

At a system level, diagnostic reliability is shaped by the two stabilisation mechanisms defined in Section 2.3: interpretive stabilisation during image reading and reporting, and confirmatory stabilisation during diagnostic confirmation (particularly image-guided biopsy and navigational procedures). In the sections that follow, we apply this framework to interpret how AI-enabled reading pathways and robotics-enabled confirmation capability contribute to variance reduction under real-world workflow and governance constraints.

9.1.1 Regulatory and governance signals Across Europe, AI and robotics used for cancer diagnosis are primarily governed as medical devices through the EU Medical Device Regulation (MDR), which requires clinical evidence proportional to risk, conformity assessment, and post-market surveillance [47]. For AI-enabled imaging tools, MDR-aligned governance is complemented by institutional quality assurance processes that monitor real-world performance, manage software updates, and document human oversight. These governance layers are not evidence of clinical effectiveness; rather, they are enabling conditions that determine whether local performance is visible, auditable, and correctable.

9.1.2 Evidence from screening workflows and human-AI interaction European evidence has increasingly shifted from algorithm-centric evaluations to workflow-centred trials and implementations. A nationwide real-world implementation study reported population-level

deployment signals for AI-supported mammography screening, illustrating how performance metrics, recall decisions, and workload can change under operational constraints [38]. In the MASAI randomised trial programme, AI-supported screen reading was evaluated within an organised screening setting with safety monitoring and workload measurement [39]. A complementary population-based trial report further characterises screening performance and subgroup patterns under AI-supported configurations [40]. Together, these studies illustrate that diagnostic impact is shaped by operational rules (who reads which cases, how disagreements are resolved, and how AI outputs are communicated) as much as by standalone algorithmic accuracy.

These studies reinforce a central adoption lesson: AI influence is mediated by human-AI interaction and the ordering of decisions. Human factors and deployment evidence show that hidden stratification and related failure modes can produce clinically meaningful errors that are not apparent in aggregate performance reporting, reinforcing the need for subgroup-aware monitoring and clear escalation rules [42].

9.1.3 Robotics as diagnostic confirmation infrastructure (Europe) Robotics adoption in Europe within diagnostic imaging is characterised by procedure-specific, centre-based deployment, reflecting both regulatory caution and the inherently local nature of interventional diagnostics. Unlike population-scale screening AI, which is evaluated across millions of examinations, the diagnostic value of robotics is assessed at the level of individual procedures where imaging findings must be translated into tissue confirmation. As a result, quantitative evidence for robotics in diagnosis is appropriately concentrated on metrics such as targeting accuracy, number of needle adjustments, and successful lesion localisation, rather than population-level outcome shifts.

In the context of CT-guided and navigational biopsy, European experience with CT-guided robotic systems provides a clear illustration of how robotics contributes to diagnostic reliability. Published evaluations report improved needle positioning success and accuracy, together with fewer needle position adjustments, when robotic assistance is used compared with freehand CT-guided approaches [33,36]. Implementation reports from European centres describe practical pathways for integrating CT-guided biopsy robotics into interventional oncology workflows, including staff training, planning protocols, and governance arrangements that enable reproducible sampling of imaging-detected lesions [43]. Reviews of biopsy and needle-robot systems similarly describe how image-guided robotic assistance can standardise trajectories and reduce operator-driven procedural variance in CT- or CBCT-guided procedures [33,36].

A similar evidentiary pattern is emerging for robotic bronchoscopy as a confirmation technology for small, peripheral pulmonary nodules that are difficult to reach with conventional bronchoscopic techniques. Early prospective European studies of shape-sensing robotic-assisted bronchoscopy report high localisation success and clinically relevant diagnostic yield in cohorts selected for challenging peripheral targets, with safety profiles that support further staged evaluation [53].

Taken together, European robotics adoption demonstrates a distinct quantitative profile: procedure volumes are smaller than screening cohorts but diagnostically sharper. Each robotics-assisted intervention is tightly coupled to an imaging-identified abnormality and directly influences diagnostic certainty at the confirmation stage. This reinforces the manuscript's central thesis

that robotics improves cancer diagnosis by reducing procedural variance at the point of tissue confirmation, complementing the role of AI in reducing interpretive variance earlier in the pathway.

To situate these procedure-level findings within the broader European enabling environment, Table 4 summarises publicly visible ecosystem signals relevant to governed adoption of imaging AI and diagnostic-confirmation robotics, explicitly distinguishing regulatory/registry signals, peer-reviewed procedure evidence, and vendor-reported deployment statements. The table is used for infrastructure mapping and implementation context rather than comparative effectiveness claims.

Table 4. Europe: implementation signals for AI-enabled imaging and diagnostic robotics.

Actor	Ecosystem role	Public signal	Diagnostic relevance	Interpretation	Source
EU MDR / EUDAMED	Regulatory and post-market surveillance backbone	MDR mandates lifecycle clinical evidence and post-market surveillance; EUDAMED supports transparency and device lifecycle traceability.	Enables governed diagnostic deployment (auditability, change control).	Infrastructure signal, not performance evidence.	EU MDR and EUDAMED overview
Radiology Health AI Register	Market visibility of CE-marked radiology AI	More than 200 CE-marked radiology AI products were listed in public European market summaries at the time of access; this should be interpreted as a registry-based market-availability signal rather than a stable denominator of clinically adopted tools.	Indicates breadth of regulated supply; supports adoption planning.	Availability signal only; counts are registry-based and time-sensitive, and do not by themselves indicate clinical validation depth, procurement, or routine uptake.	Health AI Register and secondary synthesis
BreastSlimView (Hera-Mi)	Breast imaging AI (example tool)	CE mark obtained Nov 2019; marketed in Europe.	Illustrative example for workflow / human–AI interaction.	COI-managed entry: do not use vendor-reported deployment as evidence; interpret through peer-reviewed workflow paper and Box 1 rules.	CE certification and product page
Lunit	AI (screening and chest)	Public statements of CE-MDR certification for chest imaging tool version(s).	Interpretation support in chest imaging contexts.	Treat deployment scale as vendor-reported unless independently documented.	Company announcement / trade coverage

Actor	Ecosystem role	Public signal	Diagnostic relevance	Interpretation	Source
Vara	AI mammography second read	Public reporting of CE mark under MDR; company states platform penetration in Germany.	Workflow design signal for second-reader configurations.	Keep as programme-model signal; avoid using company penetration as outcome evidence.	Company press + independent radiology trade coverage
deepc	AI orchestration platform	Platform description: single integration to manage multiple AI tools; marketplace counts are vendor-reported.	Enables scale, monitoring, and oversight across tools.	Adoption enabler; do not interpret marketplace counts as clinical uptake.	Company product documentation
Perfint MAXIO	CT-guided biopsy robotics	Peer-reviewed reports include fewer needle adjustments and accuracy or placement metrics in CT-guided robotic IO workflows.	Confirmation-stage variance reduction proxy.	Strong confirmation-infrastructure evidence (procedure metrics).	Peer-reviewed clinical series / evaluations
Medical Templates AG (Puncture Cube / navigation)	CT-guided navigation device / alternative guidance	Comparative phantom and feasibility work reported in peer-reviewed literature.	Alternative confirmation stabilisation approach.	Highlights cost and complexity trade-offs versus robotic systems.	Peer-reviewed comparative study
Intuitive – Ion (robotic bronchoscopy)	Robotic bronchoscopy for peripheral nodule confirmation	Peer-reviewed first European prospective series; CE mark timing referenced.	Lung nodule confirmation pathway relevance.	Centre-based adoption; interpret via yield, safety, and workflow feasibility (procedure evidence).	Peer-reviewed European series

Note: This table provides a non-exhaustive, illustrative set of publicly visible ecosystem signals selected to support the paper's infrastructure argument, spanning regulatory/registry visibility, workflow readiness and integration enablers, peer-reviewed procedural evidence, and publicly reported certifications. The intent is to highlight the system capabilities required to scale safely (e.g., integration, monitoring, auditability, and controlled adaptation) rather than to catalogue the market or compare products. Signals are used only to characterise implementation context and governance readiness and are not interpreted as comparative effectiveness evidence. Vendor-reported deployment statements are treated as adoption signals only and interpreted conservatively.

† **COI-managed:** Entry included to illustrate application of the framework. Claims are limited to publicly available documentation and interpreted using the evidence-handling rules in Box 1

(Section 2.5); vendor statements are treated as adoption signals unless supported by peer-reviewed outcomes.

9.2. Europe Zoom-In Case: Conflict-Managed Illustrative Example of Workflow Interaction (Hera-Mi minimised)

A peer-reviewed open-access study provides an illustrative example of how breast imaging AI can shape second-reader behaviour under double-reading conditions, without reliance on unpublished internal reports. Sossavi et al. analysed human-AI interaction in a cancer-enriched double-reading breast screening cohort, reporting diagnostic accuracy and characterising second-reader behaviour under an AI-supported configuration [50]. The value of this work for the present manuscript is its workflow-level framing: it enables discussion of how AI placement in the reading pathway can influence attention allocation, disagreement handling, and the stabilisation (or destabilisation) of interpretive decisions.

This conflict-managed example is included for illustrative purposes only and is interpreted strictly through peer-reviewed evidence and publicly available documentation under the evidence-handling rules defined in Box 1 (Section 2.5).

Public product documentation is cited only for background on intended use and device context [49].

9.3. Saudi Arabia: imaging-governance-led adoption signals in a centralised health system

Saudi Arabia is used in this analysis as an anchor case for the Gulf region because it provides a comparatively well-documented public example of system-level adoption signals for advanced diagnostic technologies under a unified policy framework [26,27,29,30]. The adoption model is shaped by centralised health-system reform, sustained investment in tertiary and quaternary care, regulator-led digital health governance, and platform-oriented scaling of diagnostic services under Vision 2030 and the Health Sector Transformation Program (HSTP) [26,27,30]. Similar dynamics are observable across neighbouring Gulf systems; however, Saudi Arabia offers particularly visible public signals of how imaging services and digital governance are being positioned within national diagnostic infrastructure under these programmes [26,27].

In this manuscript, the Saudi case is interpreted as emerging adoption signals and governance capacity, not as established comparative effectiveness. Accordingly, the analysis focuses on conditions most directly coupled to the diagnostic pathway, especially the confirmation sequence (target → sample → verify)—including: (i) imaging infrastructure and workflow integration (PACS/RIS, referral pathways, teleradiology and virtual-care scaling), (ii) regulator-led oversight for AI-enabled device software, and (iii) tertiary-centre capability to translate imaging suspicion into reproducible sampling and verification [26,27,29,30].

9.3.1 Analytic focus and evidence boundaries The Saudi case study is intentionally framed around infrastructure and governance signals rather than clinical outcomes. Public reporting describes service activation, capacity expansion, and technology partnerships but

provides limited peer-reviewed national oncology outcome evidence; accordingly, claims are kept proportionate and focused on pathway-coupled readiness conditions.

9.3.2 Imaging platform capacity, workflow integration, and governance signals From an imaging-system perspective, Saudi Arabia's transformation agenda emphasises national-scale service integration and digitally enabled care models [26,27,29]. In diagnostic imaging, platform-oriented integration is relevant because high-volume inference, controlled software update pathways, and performance monitoring depend on auditable workflows, stable connectivity between imaging systems (PACS/RIS), and clear escalation rules when AI outputs and clinician judgement diverge [30].

Two infrastructure signals are particularly relevant to imaging-based diagnostic pathways. First, the expansion of teleradiology and its operational challenges are documented in peer-reviewed Saudi analyses, including work on implementation barriers, reporting patterns, workforce capacity, and governance considerations; together, these studies highlight dependencies such as interoperability, workforce configuration, governance, and quality assurance, conditions that also shape the safe deployment of imaging AI at scale [55,56,59]. Second, national virtual-care scaling provides a system-level signal of centralised service orchestration. Seha Virtual Hospital is positioned publicly as a national virtual-care platform [28], and a peer-reviewed retrospective analysis reported rapid growth in service utilisation and increasing oncology-related case volumes over 2022–2024, supporting the interpretation of a centrally governed platform as an adoption enabler for specialist services, including imaging-enabled pathways [57].

Regulatory oversight for AI-enabled diagnostic tools is anchored by the Saudi Food and Drug Authority (SFDA), which provides guidance for digital health products and sets expectations for registration, clinical justification, and post-market oversight for software-enabled medical device functions [30]. For this manuscript's purposes, the key point is that such governance mechanisms, together with centralised service operators and tertiary centres, create enabling conditions for monitored deployment (including change control and incident review), even when peer-reviewed outcome evidence is still emerging [26,27,29,30].

9.3.3 Robotics as diagnostic confirmation infrastructure (not surgery alone) From a diagnostic standpoint, robotics in Saudi Arabia should be understood as confirmation-pathway–adjacent procedural infrastructure, not as "surgery volume" technology in isolation. In oncology workflows, diagnostic certainty often depends on translating imaging suspicion into precise, reproducible, image-guided access for tissue sampling or staging. Variability at this stage, particularly in targeting and trajectory execution, represents a clinically meaningful source of diagnostic uncertainty that image-guided robotic assistance is positioned to reduce through standardised planning and execution. Foundational navigation and image-guided intervention literature describes enabling technologies that support this role (registration, guidance, and reproducible targeting under imaging) [25]. Evidence from image-guided robotics and CT-guided intervention literature reports improved procedural precision and reduced manipulation/repositioning in CT-guided workflows, supporting the plausibility of robotics as a confirmation-stage variance-reduction mechanism when integrated with imaging pathways [33,35,36,43].

Major tertiary systems operate at a scale that enables tighter integration between advanced imaging platforms, multidisciplinary oncology pathways, and robotics-enabled procedural capability. As noted above, publicly reported robotic procedure counts are treated only as programme maturity context and not as evidence of diagnostic yield or diagnostic readiness [44–46,58]. This framing aligns with the Saudi adoption pattern: robotics platforms are embedded within centres that already possess advanced imaging capability, enterprise clinical systems, and regulator-aligned digital health governance, creating conditions in which confirmation-pathway applications can be developed as protocols mature and outcome-specific evidence accumulates [26,27,30].

Table 5 summarises the key Saudi system actors and publicly visible signals used in this case study and makes explicit how each signal is interpreted (infrastructure and governance context, not effectiveness).

Table 5. Saudi Arabia: key players and signals relevant to imaging-based cancer diagnostic confirmation (infrastructure framing)

Actor (KSA)	System role	Diagnostic relevance	Public signal	Interpretation basis
King Faisal Specialist Hospital & Research Centre (KFSHRC)	National tertiary/quaternary oncology centre; advanced imaging + robotics capability.	High-readiness context for confirmation workflows (image-guided access + procedural QA capacity) following suspicious imaging findings.	Public reporting of robotics programme activity and referral role [44,45].	Programme volumes interpreted only as maturity context (training/QA/credentialing), consistent with robotics learning-curve and governance evidence [58].
Johns Hopkins Aramco Healthcare (JHAH)	Advanced imaging + robotics centre serving a large covered population.	Supports reproducible minimally invasive access and procedural standardisation relevant to confirmation pathways.	Public reporting of programme volumes/trajectory [46].	Same interpretive constraint: counts reflect programme maturity context only; learning-curve/governance evidence supports cautious use without inferring diagnostic effectiveness [58].
Health Holdings Company (HHC)	National health services operator / network integrator.	Enables standardised workflows, referral pathways, and scalable imaging services (front-end diagnostic infrastructure).	Publicly reported partnership enabling an AI deployment model at system level [27].	Interpreted as a system-level scaling mechanism; effectiveness depends on workflow governance and monitoring [30,55,56].

Actor (KSA)	System role	Diagnostic relevance	Public signal	Interpretation basis
Seha Virtual Hospital (Ministry of Health)	National virtual hospital platform.	Signal of centrally governed scaling for specialist services (including imaging-enabled pathways).	Public platform positioning [28].	Peer-reviewed retrospective analysis documents rapid utilisation growth and oncology case-volume increases (2022–2024), supporting the "platform scaling" interpretation [57].
RapidAI	Imaging AI triage/prioritisation platform (stroke-centred deployments).	Illustrates how AI triage/routing can function as front-end infrastructure (model transferable to oncology routing logic).	Public partnership with HHC [27].	Included as a deployment-model signal; applicability to oncology depends on pathway design and evaluation [30,55,56].
SFDA (Saudi Food and Drug Authority)	Regulator for medical devices and digital health.	Governance framework for imaging AI and robotics used in diagnosis.	Digital health / software medical-device guidance [30].	Regulator guidance supports safe scaling through registration expectations, post-market oversight, and lifecycle governance [30].

Interpretation note: Entries are framed as signals of system readiness and infrastructure maturity for imaging-based diagnostic pathways. Robotics is included only where it plausibly contributes to confirmation-pathway capability (reproducible, image-guided access after imaging findings). Publicly reported robotic procedure counts are not treated as evidence of diagnostic yield improvement and are not used as proxies for diagnostic imaging readiness; they are included only as programme maturity context, consistent with peer-reviewed evidence on robotics learning curves and governance requirements [58].

9.3.4 Implications, limits, and transferability to Gulf diagnostic pathways Taken together, the Saudi case illustrates an adoption pattern in which governance capacity, platform integration, and service coordination may precede mature peer-reviewed outcome evidence. National transformation programmes shape service integration [26,27], regulator guidance defines expectations for software-enabled medical devices [30], and centrally scaled models, including teleradiology and virtual care growth, provide operational pathways through which imaging services and potentially AI-enabled workflow components can be deployed with monitoring and oversight [55–57].

At the same time, the evidentiary limits of this case are important. Public documentation provides relatively strong visibility into institutional capacity, governance arrangements, and implementation activity, but more limited peer-reviewed oncology-specific outcome evidence for many publicly described deployments. Accordingly, the Saudi case should be interpreted

primarily as an infrastructure and governance case, not as comparative evidence that technology availability alone improves diagnostic yield, accuracy, or patient outcomes.

Although core identification relied on structured searches of major bibliographic databases, supplementary searching of Google Scholar, arXiv, citation chasing, and public institutional or regulatory sources was necessarily more targeted than fully systematic. These supplementary sources were used to contextualise infrastructure, governance, and deployment signals, but they may not have captured all relevant grey literature or emerging implementation reports.

For Gulf diagnostic pathways more broadly, the main transferable lesson is therefore conditional rather than causal: where imaging platforms, referral pathways, regulatory oversight, and tertiary confirmation capacity are aligned, the organisational conditions for safer deployment of AI-enabled interpretation and robotics-supported confirmation may be stronger. However, transferability should not be assumed without local evaluation, transparent monitoring, and outcome-linked validation.

9.4. Cross-regional synthesis: what transfers, what does not

Across Europe and Saudi Arabia, a consistent pattern emerges: diagnostic improvement requires alignment between technical capability, clinical workflow, and governance. European evidence illustrates how AI can be integrated into organised screening and diagnostic workflows when configuration choices (triage vs second reader vs decision support), quality assurance, and monitoring are explicit and empirically evaluated [38–40,42]. In parallel, European robotics is evaluated through procedure-level confirmation metrics—precision, localisation success, and reduced procedural variance, appropriate to interventional diagnostics rather than population-scale endpoints [33–36,43].

The Saudi Arabia case, by contrast, is interpreted as system-readiness and governance signals, not comparative effectiveness. Public documentation indicates a centralised adoption model shaped by national transformation programmes, platform-oriented service scaling (including virtual care and teleradiology capacity), and regulator-led governance for software-enabled medical devices [26,27,30,55–57]. These features may support deployment feasibility and monitoring capacity, but peer-reviewed oncology outcome evidence remains limited for many publicly described initiatives. Accordingly, the Saudi case is used to illustrate how imaging pathways may be positioned within a unified policy framework and how confirmation-stage capability can be embedded within tertiary centres, rather than to infer diagnostic yield improvements from technology availability alone. Table 5 summarises these infrastructure signals and their interpretive limits.

Not all elements are transferable across settings. Europe's incremental, centre-by-centre evaluation model supports careful local validation and workflow optimisation but can slow scaling and create heterogeneity across sites. Saudi Arabia's centralised platform model can accelerate standardisation and oversight capacity, but its evidentiary strength depends on the availability of transparent evaluation and post-deployment monitoring outcomes. In both settings, the core lesson is that diagnostic improvement depends on aligning interpretive stabilisation and confirmatory stabilisation within governed pathways supported by auditable workflows, lifecycle governance, and explicit safety and performance monitoring.

10. Discussion

This study advances the literature by integrating artificial intelligence and robotics within a single diagnostic-systems framework rather than evaluating them as isolated technical domains. Whereas most prior reviews focus either on algorithmic performance in imaging or on technical advances in robotic procedures, the present review shows that diagnostic improvement emerges through coordinated reduction of uncertainty at distinct but interdependent stages of the diagnostic pathway.

A central conceptual contribution of this analysis is the distinction between interpretive stabilisation and confirmatory stabilisation. Artificial intelligence primarily operates upstream in the diagnostic process, where uncertainty arises from cognitive load, perceptual complexity, and inter-reader variability. In high-volume and low-prevalence settings such as population-based screening, AI may deliver clinical value by stabilising interpretation under sustained workload pressure, improving consistency across readers, and accelerating access to specialist review. Importantly, these effects are often modest at the level of individual cases but accumulate at scale, producing system-level improvements that are not adequately captured by conventional accuracy metrics alone.

Robotics, in contrast, addresses a downstream source of diagnostic uncertainty rooted in procedural variability. When imaging findings require tissue confirmation or minimally invasive access, diagnostic reliability depends on precise targeting, reproducibility, and procedural control. The evidence reviewed here indicates that robotics improves diagnostic certainty by reducing variability in needle placement, navigational access, and sampling execution, particularly in anatomically challenging cases. These gains are inherently procedure-specific and local, which explains both the centre-based nature of robotics adoption and the emphasis on precision-focused outcome measures rather than population-level performance indicators.

Crucially, this analysis demonstrates that artificial intelligence and robotics are complementary rather than substitutive. AI cannot resolve uncertainty introduced during diagnostic confirmation, just as robotics cannot compensate for upstream interpretive inconsistency. Diagnostic systems that deploy one technology without the other risk redistributing uncertainty rather than resolving it. By contrast, when interpretive stabilisation and confirmatory stabilisation are aligned within a coherent workflow, the cumulative effect is a reduction in repeat procedures, diagnostic delay, and downstream decision uncertainty.

The comparative case studies further underscore the importance of institutional context in shaping adoption trajectories. In Europe, regulated and incremental deployment of imaging AI reflects the constraints and strengths of population-scale screening programmes, where governance, accountability, and clinician trust are central. In the Gulf region, particularly in Saudi Arabia, adoption has been driven by system-level readiness, with advanced imaging platforms, enterprise-scale AI deployment, and robotics embedded within tertiary centres designed to manage complex oncology pathways. Despite differences in governance models and pace of implementation, both regions converge on a shared diagnostic logic: sustainable improvement arises when variability is reduced at both interpretive and confirmatory stages through integrated, governed systems.

10.1. Limitations

This work is an integrative, infrastructure-oriented synthesis and therefore has limitations relative to formal systematic reviews (Sections 2.1–2.6). The included evidence is heterogeneous across modalities, cancers, and study designs. Accordingly, we present a qualitative synthesis rather than pooled effect estimates (Sections 2.1–2.4). Quality appraisal tools (e.g., QUADAS-2/PROBAST/CLAIM) were used to weight interpretation (Section 2.4), and we report these assessments narratively rather than as a single quantitative score; studies were not included or excluded solely on checklist completion (Section 2.4).

Case-study analyses rely on publicly available regulatory and institutional materials, which can be incomplete, selectively reported, or lag operational reality (Section 2.5). Accordingly, the case studies are used to illustrate governance and scaling conditions rather than to assert comparative national performance (Sections 2.5 and 9.4). For Saudi Arabia specifically, we avoid using surgical robotic procedure volumes as a proxy for diagnostic imaging readiness and instead focus on imaging- and governance-relevant signals; nevertheless, any country-level inference remains indirect and should be interpreted cautiously (Sections 2.5 and 9.3.4).

Finally, the illustrative industry–academic case (Hera-Mi) warrants explicit transparency because one author is employed by the company (Section 2.5 and Declarations/Competing interests). To manage potential bias, we limited the case to publicly available documentation and avoided proprietary performance claims (Section 2.5 and Declarations/Competing interests). The case is therefore intended to illustrate application of the framework rather than to evaluate a specific product (Section 2.5).

Across all technological layers, the findings reinforce that governance is not merely a constraint on innovation but a determinant of diagnostic performance (Sections 5–8). Bias, performance drift, procedural variability, and accountability gaps undermine diagnostic value when left unmanaged (Sections 4 and 5–8). Health systems that treat AI and robotics as safety-critical diagnostic infrastructure, subject to continuous monitoring, audit, and role clarity, are better positioned to achieve durable diagnostic gains than those pursuing fragmented or pilot-driven deployment strategies (Sections 6–9).

Rather than framing progress in cancer diagnosis as a competition between algorithms and clinicians or between robotic and manual procedures, this analysis reframes the problem as one of system design (Sections 3–5). Diagnostic improvement depends on how technologies are integrated into workflows, aligned with professional responsibility, and governed over time (Sections 4–9). Evaluated through this lens, artificial intelligence and robotics emerge not as isolated innovations but as coordinated instruments in the ongoing transformation of cancer diagnosis into resilient, accountable clinical infrastructure (Sections 9–11).

11. Conclusion

Artificial intelligence and robotics function as complementary pillars of modern cancer diagnosis, addressing distinct but interdependent sources of diagnostic uncertainty. By distinguishing interpretive stabilisation from confirmatory stabilisation, this study reframes diagnostic value away from isolated performance metrics toward variance reduction at the system level. The findings demonstrate that durable improvement in cancer diagnosis emerges when AI and robotics are integrated into governed, end-to-end diagnostic infrastructure rather than deployed as standalone tools. This infrastructure-oriented perspective provides a foundation for future clinical deployment, regulatory alignment, and research on emerging diagnostic technologies, including micro- and nanorobotics.

Declarations

Competing interests. M. Tardy is employed by Hera-Mi and serves as Chief Scientific Officer. He is also a co-author of reference [50], which is included in this manuscript as a conflict-managed illustrative workflow example. Claims drawn from this reference are restricted to its published peer-reviewed content; no proprietary performance datasets, unpublished dashboards, or internal reports were used as evidence. K. Chebib independently reviewed the inclusion and interpretation of this reference. K. Chebib declares no additional competing interests. The paper's conclusions are based on synthesis of peer-reviewed literature and public sources.

Author contributions. K. Chebib: Conceptualisation, Methodology, Investigation, Data curation, Formal analysis, Writing – original draft, Visualisation, and independent appraisal and interpretation of the evidence base. M. Tardy: Validation limited to technical/regulatory description of the conflict-managed illustrative case, and Writing (review & editing). Both authors: Critical revision for important intellectual content and approval of the final manuscript.

Role of the partnering organisation (Hera-Mi). Hera-Mi did not provide funding for this study and did not influence evidence selection, appraisal, or the interpretation of peer-reviewed findings. The company's involvement was limited to factual clarification of publicly available product/regulatory information included in the illustrative case. The corresponding author retained full editorial control and responsibility for the manuscript.

Materials availability. The evidence-map extraction template is provided in the Supplementary material to support reproducibility. No proprietary datasets were used in this review.

Data availability. All sources used for this study are publicly available and cited in the reference list. No new patient-level datasets were generated or analysed.

Funding. No external funding was received for this study.

References

[1] Aggarwal R, Sounderajah V, Martin G, et al. Diagnostic accuracy of deep learning in medical imaging: a systematic review and meta-analysis. NPJ Digit Med. 2021;4:65. doi:10.1038/s41746-021-00429-0.

[2] Bi WL, Hosny A, Schabath MB, et al. Artificial intelligence in cancer imaging: clinical challenges and applications. CA Cancer J Clin. 2019;69(2):127-157. doi:10.3322/caac.21552.

[3] Coppola F, Faggioni L, Regge D, et al. Artificial intelligence: radiologists' expectations and opinions gleaned from a nationwide online survey. Radiol Med. 2021;126(1):63-71. doi:10.1007/s11547-020-01240-x.

[4] Higaki T, Nakamura Y, Tatsugami F, Nakaura T, Awai K. Improvement of image quality at CT and MRI using deep learning. Jpn J Radiol. 2019;37(1):73-80. doi:10.1007/s11604-018-0793-2.

[5] Hosny A, Parmar C, Quackenbush J, Schwartz LH, Aerts HJWL. Artificial intelligence in radiology. Nat Rev Cancer. 2018;18(8):500-510. doi:10.1038/s41568-018-0016-5.

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

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

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

[9] Sendak MP, D'Arcy J, Kashyap S, et al. A path for translation of machine learning products into healthcare delivery. EMJ Innov. 2020. doi:10.33590/emjinnov/19-00172.

[10] Wu E, Wu K, Daneshjou R, Ouyang D, Ho DE, Zou J. How medical AI devices are evaluated: limitations and recommendations from an analysis of FDA approvals. Nat Med. 2021;27(4):582-584. doi:10.1038/s41591-021-01312-x.

[11] van Leeuwen KG, Schalekamp S, Rutten MJCM, et al. Artificial intelligence in radiology: 100 commercially available products and their scientific evidence. Eur Radiol. 2021;31(7):3797-3804. doi:10.1007/s00330-020-07569-z.

[12] Antonissen N, Tryfonos O, Houben IB, Jacobs C, de Rooij M, van Leeuwen KG. Artificial intelligence in radiology: 173 commercially available products and their scientific evidence. Eur Radiol. 2023;33(4):2376–2387. doi:10.1007/s00330-022-09186-5.

[13] Collins GS, Moons KGM. Reporting of artificial intelligence prediction models. Lancet. 2019;393(10181):1577-1579. doi:10.1016/S0140-6736(19)30072-6.

[14] Liu X, Rivera SC, Moher D, et al. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. Nat Med. 2020;26(9):1364-1374. doi:10.1038/s41591-020-1034-x.

[15] Vasey B, Nagendran M, Campbell B, et al. Reporting guideline for the early-stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI. Nat Med. 2022;28(5):924-933. doi:10.1038/s41591-022-01772-3.

[16] Lauritzen AD, Haldrup M, Petersen NØ, et al. Deep learning-based detection of lung cancer in low-dose CT: a prospective study. Lancet Digit Health. 2022;4(6):e415-e423. doi:10.1016/S2589-7500(22)00073-7.

[17] Ardila D, Kiraly AP, Bharadwaj S, et al. End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. Nat Med. 2019;25(6):954-961. doi:10.1038/s41591-019-0447-x.

[18] Shelmerdine SC, Arthurs OJ. Deep learning in radiology: a review of the current status. Clin Radiol. 2020;75(5):354-362. doi:10.1016/j.crad.2019.12.005.

[19] O'Neill TJ, Xi Y, Stehel E, et al. Active reprioritization of the reading worklist using artificial intelligence has a beneficial effect on the turnaround time for interpretation of head CT with intracranial hemorrhage. Radiol Artif Intell. 2020;3(2):e200024. doi:10.1148/ryai.2020200024.

[20] Benjamens S, Dhunnoo P, Meskó B. The state of artificial intelligence-based FDA-approved medical devices and algorithms: an online database. NPJ Digit Med. 2020;3:118. doi:10.1038/s41746-020-00324-0.

[21] European Commission. Ethics guidelines for trustworthy AI [Internet]. Brussels: EC; 2019. Available from: https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai. Accessed 28 January 2026.

[22] World Health Organization. Ethics and governance of artificial intelligence for health [Internet]. Geneva: WHO; 2021. Available from: https://www.who.int/publications/i/item/9789240029200. Accessed 28 January 2026.

[23] International Organization for Standardization. ISO 13485:2016 Medical devices — Quality management systems — Requirements for regulatory purposes. Geneva: ISO; 2016.

[24] International Organization for Standardization. ISO/IEC 23894:2023 Artificial intelligence — Guidance on risk management. Geneva: ISO; 2023.

[25] National Institute for Health and Care Excellence. Evidence standards framework for digital health technologies [Internet]. London: NICE; 2022. Available from: https://www.nice.org.uk/what-nice-does/digital-health/evidence-standards-framework-esf-for-digital-health-technologies. Accessed 2 February 2026.

[26] Saudi Vision 2030. Health Sector Transformation Program [Internet]. Available from: https://www.vision2030.gov.sa/en/explore/programs/health-sector-transformation-program. Accessed 2 February 2026.

[27] Ministry of Health (Saudi Arabia). Health Sector Transformation Program [Internet]. Riyadh: MoH; 2026. Available from: https://www.hstp.gov.sa/en. Accessed 2 February 2026.

[28] Seha Virtual Hospital. Seha Virtual Hospital [Internet]. Riyadh: Ministry of Health; 2026. Available from: https://svh.moh.gov.sa. Accessed 2 February 2026.

[29] Saudi Ministry of Health. Seha Virtual Hospital [Internet]. MOH Announcements. 28 Feb 2022. Available from: https://www.moh.gov.sa/en/Ministry/MediaCenter/Ads/Pages/Ads-2022-02-28-001.aspx. Accessed 2 February 2026.

[30] Saudi Food and Drug Authority (SFDA). Guidance on Artificial Intelligence (AI) and Machine Learning (ML) Technologies Based Medical Devices (MDS-G-010) [Internet]. Riyadh: SFDA; 2023. Available from: https://www.sfda.gov.sa/sites/default/files/2023-01/MDS-G010ML.pdf. Accessed 2 February 2026.

[31] Cleary K, Clifford M. Needle and biopsy robots: clinical applications and technical design. Cardiovasc Intervent Radiol. 2021;44(1):1-15. doi:10.1007/s00270-020-02634-1.

[32] Ergeneman O, Abbott JJ, Nelson BJ. Micro- and nanorobots for biomedical applications. Annu Rev Biomed Eng. 2021;23:177-202. doi:10.1146/annurev-bioeng-082120-021242.

[33] Kettenbach J, Kronreif G, Figl M, et al. Robot-assisted CT-guided percutaneous needle placement: evaluation of accuracy and efficiency. Cardiovasc Intervent Radiol. 2022;45(2):165-173. doi:10.1007/s00270-021-03014-0.

[34] Chen J, Huang J, Zhang Y, et al. Robotic bronchoscopy for peripheral pulmonary lesions: a systematic review and meta-analysis. Chest. 2023;164(2):441-452. doi:10.1016/j.chest.2023.03.011.

[35] Chen AC, Pastis NJ, Mahajan AK, et al. Robotic bronchoscopy for pulmonary lesions: a multicenter pilot and feasibility study (BENEFIT). Chest. 2021;159(1):369-377. doi:10.1016/j.chest.2020.08.2047.

[36] Chiu HY, Lin CY, Chen JY, et al. Robotic-assisted CT-guided percutaneous intervention: accuracy and feasibility. Eur Radiol. 2020;30(2):1003-1010. doi:10.1007/s00330-019-06424-8.

[37] De Fauw J, Ledsam JR, Romera-Paredes B, et al. Clinically applicable deep learning for diagnosis and referral in retinal disease. Nat Med. 2018;24(9):1342-1350. doi:10.1038/s41591-018-0107-6.

[38] Eisemann N, Bunk S, Mukama T, et al. Nationwide real-world implementation of AI for cancer detection in population-based mammography screening. Nat Med. 2025;31(3):917-924. Epub 7 Jan 2025. doi:10.1038/s41591-024-03408-6.

[39] Lång K, Josefsson V, Larsson A-M, et al. Artificial intelligence-supported screen reading versus standard double reading in the Mammography Screening with Artificial Intelligence trial (MASAI): a clinical safety analysis of a randomised, controlled, non-inferiority, single-blinded, screening accuracy study. Lancet Oncol. 2023;24(8):936-944. doi:10.1016/S1470-2045(23)00298-X.

[40] Hernström V, Josefsson V, Sartor H, et al. Screening performance and characteristics of breast cancer detected in the Mammography Screening with Artificial Intelligence trial (MASAI): a randomised, controlled, parallel-group, non-inferiority, single-blinded, screening accuracy study. Lancet Digit Health. 2025;7(3):e175-e183. Epub 3 Feb 2025. doi:10.1016/S2589-7500(24)00267-X.

[41] Houssami N, Marinovich ML. AI for mammography: making double screen-reading history. Lancet Digit Health. 2025;7(3):e168-e169. doi:10.1016/j.landig.2025.01.004.

[42] Oakden-Rayner L, Dunnmon J, Carneiro G, Ré C. Hidden stratification causes clinically meaningful failures in machine learning for medical imaging. Proc ACM Conf Health Inference Learn (CHIL). 2020:151–159. doi:10.1145/3368555.3384468.

[43] Johnston EW, Basso J, Winfield J, McCall J, Khan N, Messiou C, Koh D-M, Fotiadis N. Starting CT-guided robotic interventional oncology at a UK centre. Br J Radiol. 2022;95(1134):20220217. doi:10.1259/bjr.20220217.

[44] King Faisal Specialist Hospital & Research Centre. Annual Report 2023 [Internet]. Riyadh: KFSHRC; 2024. Available from: https://www.kfshrc.edu.sa/-/media/projects/kfshrc/annual-report/2023. Accessed 2 February 2026.

[45] King Faisal Specialist Hospital & Research Centre. KFSHRC Pushes the Boundaries of Robotic Surgery with World-First Achievements [Internet]. 2025 Oct 15. Available from: https://www.kfshrc.edu.sa. Accessed 2 February 2026.

[46] Johns Hopkins Aramco Healthcare. Inside JHAH's Robotic Surgery Program: Saudi Arabia's da Vinci record-holder [Internet]. 2025 Aug 4. Available from: https://www.jhah.com/en/news/inside-jhah-robotic-surgery-program. Accessed 2 February 2026.

[47] European Parliament and Council of the European Union. Regulation (EU) 2017/745 of 5 April 2017 on medical devices (MDR) [Internet]. Off J Eur Union. 2017. Available from: https://eur-lex.europa.eu/eli/reg/2017/745/oj/eng. Accessed 28 January 2026.

[48] European Commission, Directorate-General for Health and Food Safety. Study on the deployment of AI in healthcare—Final report [Internet]. 2025 Aug 8. Available from: https://health.ec.europa.eu/publications/study-deployment-ai-healthcare-publications-office-eurep_en. Accessed 2 February 2026.

[49] Hera-MI. Breast-SlimView®: AI solution for mammography [Internet]. Available from: https://www.hera-mi.com/en/ai-solution-for-mammography/. Accessed 2 February 2026.

[50] Sossavi E, Tardy M, Hurstel F, et al. Human–AI interaction in a cancer-enriched double-reading breast screening cohort: diagnostic accuracy and second-reader behavior. Cancer Imaging. 2026;26(1):29. doi:10.1186/s40644-026-00995-0.

[51] European Parliament and Council of the European Union. Regulation (EU) 2024/1689 of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts [Internet]. Off J Eur Union. 2024. Available from: https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng. Accessed 2 February 2026.

[52] U.S. Food and Drug Administration. Marketing submission recommendations for a predetermined change control plan for artificial intelligence/machine learning (AI/ML)-enabled device software functions: guidance for industry and Food and Drug Administration staff [Internet]. Silver Spring (MD): FDA; 2025 Aug

18. Available from: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/marketing-submission-recommendations-predetermined-change-control-plan-artificial-intelligence. Accessed 2 February 2026.

[53] Brock JM, Dittrich AS, Kontogianni K, Heussel CP, Klotz LV, Winter H, Schellenberg M, Keppler U, Herth FJF. First European Results of Shape-Sensing Robotic-Assisted Bronchoscopy for Peripheral Pulmonary Lesions. Respiration. 2025;104(3):234-243. doi:10.1159/000536212.

[54] U.S. Food and Drug Administration. Artificial Intelligence-Enabled Medical Devices [Internet]. Silver Spring (MD): FDA. Available from: https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-enabled-medical-devices. Accessed 2 February 2026.

[55] Alahmad H, Almanaa MA, Abanomy A, et al. Navigating challenges in teleradiology implementation: a case study from Saudi Arabia's healthcare system. J Multidiscip Healthc. 2024;17:225-236. doi:10.2147/JMDH.S447201.

[56] Alruwaili AR, Alshammari AA, Alsalhi FM, et al. Teleradiology in Saudi Arabia: a national survey and retrospective review of associated MRI reports. BMC Health Serv Res. 2024;24(1):247. doi:10.1186/s12913-024-10632-2.

[57] Alabdulaali MK, Aldossary MS, Alsubaie MS, Zaatari ES, Alhomod AS. Digital transformation in telemedicine health care: a retrospective analysis of Seha Virtual Hospital outcomes. J Med Internet Res. 2024;26:e50511. doi:10.2196/50511.

[58] Soomro NA, Hashimoto DA, Porteous AJ, et al. Systematic review of learning curves in robot-assisted surgery. BJS Open. 2020;4(1):27-44. doi:10.1002/bjs5.50235.

[59] Al-Dahery ST, Alsharif WM, Alamri FH, Nawawi SA, Mofti WK, Alhazmi FH, Alshamrani KM, Suliman AG, Qurashi AA. The role of teleradiology during COVID-19 outbreak: Saudi radiologists' perspectives. Saudi Med J. 2023;44(2):202-210. doi:10.15537/smj.2023.44.2.20220793.

[60] International Medical Device Regulators Forum. Software as a Medical Device (SaMD): Clinical Evaluation [Internet]. IMDRF; 2017. Available from: http://www.imdrf.org/docs/imdrf/final/technical/imdrf-tech-170921-samd-n41-clinical-evaluation_1.pdf. Accessed 2 February 2026.

[61] U.S. Food and Drug Administration. Artificial Intelligence/Machine Learning (AI/ML)-Based Software as a Medical Device (SaMD) Action Plan [Internet]. Silver Spring (MD): FDA; 2021. Available from: https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-samd. Accessed 2 February 2026.

[62] European Commission. Proposal for a regulation on the European Health Data Space (EHDS) [Internet]. Brussels: EC; 2022. Available from: https://health.ec.europa.eu/ehealth-digital-health-and-care/european-health-data-space_en. Accessed 2 February 2026.

[63] Nawabi J, Eminovic S, Weissflog JS, et al. Artificial intelligence in radiology: A comparative analysis of reimbursement and regulatory developments in the US and EU. Eur J Radiol Artif Intell. 2026;5:100056. doi:10.1016/j.ejrai.2025.100056.

[64] European Commission. Medical Device Coordination Group guidance: MDCG 2020-1 (Guidance on clinical evaluation) [Internet]. Brussels: European Commission; 2020. Available from: https://health.ec.europa.eu/medical-devices-sector/guidance-documents_en. Accessed 2 February 2026.

[65] Wang Z. Artificial intelligence in Chinese healthcare: a review of applications and future prospects. *Biomed Eng Lett.* 2025;15:1065–1072. doi:10.1007/s13534-025-00515-2.

[66] Ministry of Health (Saudi Arabia). National eHealth Strategy [Internet]. Riyadh: MoH; 2020. Available from: https://www.moh.gov.sa. Accessed 2 February 2026.

[67] Department of Health and Social Care. Initial code of conduct for data-driven health and care technology: a guide to good practice [Internet]. London: DHSC; 2021. Available from: https://www.gov.uk/government/publications/code-of-conduct-for-data-driven-health-and-care-technology/initial-code-of-conduct-for-data-driven-health-and-care-technology. Accessed 28 January 2026.

[68] Medical Device Coordination Group. MDCG 2019-11 Rev.1: Qualification and classification of software under Regulation (EU) 2017/745 and Regulation (EU) 2017/746 [Internet]. Brussels: European Commission; 2025. Available from: https://health.ec.europa.eu/system/files/2025-06/mdcg_2019_11_rev1_en.pdf. Accessed 2 February 2026.

A. Supplementary tables

These tables provide illustrative, non-exhaustive examples of ecosystem actors and publicly observable adoption signals. They are descriptive and should not be interpreted as systematic evidence of effectiveness.

Table A1. Europe: illustrative ecosystem actors and adoption signals relevant to imaging-based cancer diagnosis

Actor	Role	Public signal	Diagnostic relevance	Interpretation
EU MDR / EUDAMED	Regulatory and surveillance backbone	Risk-based medical device oversight and post-market surveillance obligations.	Enables governed diagnostic deployment.	Infrastructure signal (not a performance metric).
Organised breast screening programmes	Operational pathway standardisation	Established double-reading workflows with QA metrics.	Creates a high-volume setting where AI can affect workload.	Determines where AI can be inserted and monitored.
National screening authorities / QA bodies	Accountability and audit	National audit cycles, recall thresholds, reader standards.	Maintains safety while workflows evolve.	Governance signal, not evidence of AI benefit.
Hospital PACS/RIS vendors and interoperability initiatives	Integration substrate	Routine digital imaging workflows and structured reporting.	Enables deployment and monitoring of AI and robotics.	Readiness signal for measurement and traceability.
Peer-reviewed workflow trials (e.g., MASAI programme)	Evidence generation in real practice	Trial designs that report workload, recall, safety, and subgroup patterns.	Supports inference about workflow effects.	Stronger than standalone accuracy studies.
Human factors literature on AI decision support	Sociotechnical risk management	Evidence of automation bias and reader-AI interaction effects.	Guides UI, escalation, and accountability design.	Implementation determinant.

Table A2. Saudi Arabia: imaging & robotics ecosystem adoption signals (public)

Actor (KSA)	System role	Diagnostic relevance	Public signal	Interpretation
King Faisal Specialist Hospital & Research Centre (KFSHRC)	National tertiary oncology and complex intervention hub.	High-acuity oncology confirmation pathways; interventional capacity readiness.	Public reporting of robotics programmes and volumes [44,45].	Volume is a contextual signal; not a proxy for diagnostic readiness.

Actor (KSA)	System role	Diagnostic relevance	Public signal	Interpretation
Ministry of Health (MOH)	National policy, service models, and transformation programmes.	Shapes infrastructure, PACS/RIS, referral networks, and centre designation.	Health sector transformation programme documentation [26,29].	Governance signal only; it does not establish clinical outcomes.
Saudi Food and Drug Authority (SFDA)	Medical software and device oversight.	Regulates AI-enabled imaging as medical software/devices; post-market expectations.	Digital health product guidance including AI-enabled functions [30].	Regulatory clarity enables safer deployment; effectiveness still depends on local workflow and evidence.
National digital health operators / platforms	Digital infrastructure for scale (data, connectivity, service orchestration).	Enables monitoring, audit, and integration of AI outputs into workflows.	National digital health and virtual care initiatives [29].	Readiness signal; clinical impact requires validated use cases.
Johns Hopkins Aramco Healthcare (JHAH)	High-capacity tertiary provider with public procedure reporting.	Illustrates availability of advanced robotics services and reporting norms.	Public reporting of robotic procedures [46].	Not used as a diagnostic readiness proxy; included as transparency about public reporting.
Radiology and IR service lines in tertiary centres	Confirmation-stage execution (target → sample → verify).	Determines whether imaging findings can be reliably translated into tissue diagnosis.	Service activation announcements and institutional reporting [44,45].	Treat as emerging adoption signals; avoid causal claims absent peer-reviewed outcomes.

B. Database search strategies (search strings)

Scope and limits:

Date range: January 2018 to March 2026.

Peer-reviewed literature stream only; public regulatory and institutional sources were mapped separately.

Searches were limited to English-language records where supported by database filtering.

B.1. MEDLINE (PubMed)

AI in cancer imaging (interpretation / triage / workflow)

((cancer[Title/Abstract] OR oncology[Title/Abstract] OR neoplasm*[Title/Abstract] OR tumor*[Title/Abstract] OR tumour*[Title/Abstract])

AND (imaging[Title/Abstract] OR radiology[Title/Abstract] OR mammograph*[Title/Abstract] OR tomosynthesis[Title/Abstract]

OR CT[Title/Abstract] OR "computed tomography"[Title/Abstract] OR MRI[Title/Abstract] OR "magnetic resonance"[Title/Abstract]

OR ultrasound[Title/Abstract] OR PET[Title/Abstract])

AND ("artificial intelligence"[Title/Abstract] OR "machine learning"[Title/Abstract] OR "deep learning"[Title/Abstract]

OR CAD[Title/Abstract] OR "computer-aided"[Title/Abstract] OR "computer aided"[Title/Abstract]))

Robotics / navigation for image-guided diagnostic confirmation (biopsy / bronchoscopy / targeting / sampling)

((cancer[Title/Abstract] OR oncology[Title/Abstract] OR neoplasm*[Title/Abstract] OR tumor*[Title/Abstract] OR tumour*[Title/Abstract])

AND (biopsy[Title/Abstract] OR "tissue sampling"[Title/Abstract] OR "image-guided"[Title/Abstract] OR "image guided"[Title/Abstract]

OR navigation[Title/Abstract] OR bronchoscopy[Title/Abstract] OR "robotic bronchoscopy"[Title/Abstract]

OR "CT-guided"[Title/Abstract] OR "ultrasound-guided"[Title/Abstract] OR "MRI-guided"[Title/Abstract])

AND (robot*[Title/Abstract] OR "robot-assisted"[Title/Abstract] OR "robotic-assisted"[Title/Abstract]

OR "shape-sensing"[Title/Abstract] OR "shape sensing"[Title/Abstract] OR "electromagnetic navigation"[Title/Abstract]))

Combined query (broad "AI OR robotics" capture, then screen)

((cancer[Title/Abstract] OR oncology[Title/Abstract] OR neoplasm*[Title/Abstract] OR tumor*[Title/Abstract] OR tumour*[Title/Abstract])

AND (imaging[Title/Abstract] OR radiology[Title/Abstract] OR biopsy[Title/Abstract] OR bronchoscopy[Title/Abstract] OR "image-guided"[Title/Abstract])

AND (("artificial intelligence"[Title/Abstract] OR "machine learning"[Title/Abstract] OR "deep learning"[Title/Abstract] OR CAD[Title/Abstract])

OR (robot*[Title/Abstract] OR "robot-assisted"[Title/Abstract] OR navigation[Title/Abstract])))

B.2. Scopus (TITLE-ABS-KEY)

AI stream

TITLE-ABS-KEY(

(cancer OR oncology OR neoplasm* OR tumor* OR tumour*)

AND (imaging OR radiology OR mammograph* OR tomosynthesis OR CT OR "computed tomography" OR MRI OR "magnetic resonance" OR ultrasound OR PET)

AND ("artificial intelligence" OR "machine learning" OR "deep learning" OR CAD OR "computer-aided" OR "computer aided")

)

Robotics stream

TITLE-ABS-KEY(

(cancer OR oncology OR neoplasm* OR tumor* OR tumour*)

AND (biopsy OR "tissue sampling" OR "image-guided" OR "image guided" OR navigation OR bronchoscopy OR "robotic bronchoscopy"

OR "CT-guided" OR "ultrasound-guided" OR "MRI-guided")

AND (robot* OR "robot-assisted" OR "robotic-assisted" OR "shape-sensing" OR "shape sensing" OR "electromagnetic navigation")

)

B.3. Web of Science Core Collection (TS=)

AI stream

TS=((cancer OR oncology OR neoplasm* OR tumor* OR tumour*)

AND (imaging OR radiology OR mammograph* OR tomosynthesis OR CT OR "computed tomography" OR MRI OR "magnetic resonance" OR ultrasound OR PET)

AND ("artificial intelligence" OR "machine learning" OR "deep learning" OR CAD OR "computer-aided" OR "computer aided"))

Robotics stream

TS=((cancer OR oncology OR neoplasm* OR tumor* OR tumour*)

AND (biopsy OR "tissue sampling" OR "image-guided" OR "image guided" OR navigation OR bronchoscopy OR "robotic bronchoscopy"

OR "CT-guided" OR "ultrasound-guided" OR "MRI-guided")

AND (robot* OR "robot-assisted" OR "robotic-assisted" OR "shape-sensing" OR "shape sensing" OR "electromagnetic navigation"))

B.4. IEEE Xplore

AI stream

("cancer" OR "oncology")

AND ("medical imaging" OR radiology OR mammography OR tomosynthesis OR CT OR MRI OR ultrasound OR PET)

AND ("artificial intelligence" OR "machine learning" OR "deep learning")

Robotics stream

("cancer" OR "oncology")

AND (biopsy OR "image-guided" OR "image guided" OR navigation OR bronchoscopy)

AND (robot* OR "robot-assisted" OR "robotic-assisted" OR "electromagnetic navigation" OR "shape sensing" OR "shape-sensing")

B.5. Supplementary source searches (reproducible description)

Google Scholar: used for supplementary coverage and to capture clinically oriented implementations not consistently indexed across databases. Search used the same core concepts (e.g., *cancer imaging deep learning workflow*, *robotic bronchoscopy diagnostic yield*, *CT-guided biopsy robot needle adjustments*), with date filtering to the defined review window; results were screened using the same title/abstract then full-text eligibility criteria.

arXiv: used to identify engineering advances relevant to diagnostic confirmation/robotics and ML methods; arXiv items were retained only when they supported infrastructure/governance interpretation or were later peer-reviewed (otherwise treated as supplementary context rather than peer-reviewed evidence).

Citation chasing and reference-list screening: performed forward/backward from high-relevance reviews and consensus/regulatory documents to ensure capture of influential clinical deployment studies and robotics confirmation series.