International Journal of Transformative Multidisciplinary Studies (IJTMS), 2 (2), 2026

Review Article

From Image Acquisition to Algorithmic Accountability: Reframing Research Methodology in Radiologic Technology for the Age of Artificial Intelligence

Mark Alipio [1,*]

Received: 02 February 2026; Revised: 29 May 2026;
Accepted: 06 June 2026; Published: 08 June 2026

DOI: https://doi.org/10.66074/R3B7H4D9

Abstract

Artificial intelligence has progressed from experimental image classification to tools influencing acquisition, triage, reconstruction, image quality assessment, dose optimization, workflow, education, and departmental governance. However, medical imaging literature often remains model-centered, emphasizing area under the curve, sensitivity, specificity, or reader performance, while many Radiologic Technology studies focus on perceptions, readiness, and attitudes. This narrative review reframes AI-related Radiologic Technology research as a methodological issue rather than a technology-adoption topic. It argues that radiologic technologists shape how imaging data are produced, interpreted, repeated, rejected, archived, and used for algorithmic learning. Thus, AI-era research must treat acquisition protocols, positioning, exposure parameters, dose indices, patient preparation, equipment variation, image quality, workflow behavior, trust, override decisions, and post-deployment monitoring as core design elements. Synthesizing recent AI reporting standards and evaluation frameworks, the review identifies gaps in validation, human-AI interaction, implementation, generalizability, equity, accountability, and AI literacy. It proposes the RADIATE-AI framework to guide safer, locally valid, and methodologically mature studies.

Keywords: artificial intelligence, education, radiography, radiologic technology, research

[1] Iligan Medical Center College, Philippines
* mark.alipio@imcc.edu.ph

Volume 2, Issue 2,
June 2026

1. Introduction

Radiologic Technology research has reached a point where familiar methodological habits no longer fully match the conditions of practice. The field has long examined image quality, radiation protection, patient care, clinical education, professional roles, and departmental workflow. These concerns remain central, but artificial intelligence has changed the research environment in which those concerns are studied. AI can alter how images are acquired, how protocols are selected, how quality is judged, how urgent findings are triaged, how examinations move through a department, and how staff learn to trust or challenge algorithmic advice. A manuscript that treats AI only as another diagnostic tool risks missing the deeper problem: AI changes the methods by which Radiologic Technology research should ask questions, define variables, choose outcomes, document context, and judge value.

Medical imaging AI has matured through impressive technical progress, but the research literature continues to show methodological imbalance. Many studies report high discrimination, segmentation accuracy, or reader performance, yet external validation, prospective evaluation, workflow analysis, health equity, and post-deployment surveillance remain less common than performance modeling (Liu et al., 2019; Nagendran et al., 2020; Yu et al., 2022). Reporting standards have responded to this weakness. CONSORT-AI and SPIRIT-AI extended trial reporting to clinical AI interventions, DECIDE-AI addressed early-stage clinical evaluation, TRIPOD+AI revised prediction-model reporting, STARD-AI updated diagnostic accuracy reporting for AI-enabled tests, CLAIM 2024 refined reporting for AI in medical imaging, and FUTURE-AI proposed principles for trustworthy AI in health care (Collins et al., 2024; Cruz Rivera et al., 2020; Lekadir et al., 2025; Liu et al., 2020; Sounderajah et al., 2025; Tejani et al., 2024; Vasey et al., 2022). These documents improve the general standard for AI research, but they do not automatically solve the particular methodological questions faced by radiologic technologists.

The role of the radiologic technologist is not peripheral to imaging AI. Images do not arrive in datasets as neutral artifacts. They result from patient preparation, communication, positioning, exposure selection, immobilization, breath-hold instruction, protocol adherence, equipment choice, artifact management, and local quality control. A chest radiograph, CT series, mammogram, fluoroscopic acquisition, or MRI sequence reflects technical decisions made before any algorithm evaluates pixels or metadata. A poorly positioned image can reduce diagnostic confidence; a protocol shift can alter the distribution of intensities; a portable examination can introduce artifacts and context that differ from fixed-room imaging. These are not minor technical details. In AI-enabled research, they are determinants of validity.

A methodological reset is also needed because many Radiologic Technology studies on AI remain framed as readiness, awareness, or perception studies. Such studies are useful in the early stage of a professional transition, and radiographer surveys have documented concerns about training, job security, ethical responsibility, and professional adaptation (Aldhafeeri, 2022; Botwe et al., 2021; Hardy & Harvey, 2020). The next phase of scholarship requires stronger designs. RT researchers need validation studies that record acquisition conditions, simulation studies that examine radiographer-AI collaboration, implementation studies that measure adoption and fidelity, interrupted time-series studies that test effects on repeat rates or turnaround time, educational intervention studies that assess AI literacy, and local generalizability studies that test performance across equipment, patient groups, and institutions.

This narrative review therefore advances a central argument: Radiologic Technology research in the AI era should move beyond model-performance reporting and perception surveys toward acquisition-aware, human-centered, implementation-sensitive, ethically accountable, and locally validated research designs. The review does not claim that AI will replace professional judgment. Rather, it treats AI as a socio-technical intervention embedded in imaging work. Its contribution is the RADIATE-AI framework, a practical methodological guide for future RT theses, journal manuscripts, protocol development, curriculum reform, quality assurance research, and institutional AI adoption studies.

2. Scope and Conceptual Boundaries of the Review

This review focuses on research methodology in Radiologic Technology, not on the full universe of AI applications in radiology. Radiology literature often centers on radiologist interpretation, model architecture, lesion detection, segmentation, report generation, or diagnostic performance. Those topics matter, but they do not exhaust the research questions that radiologic technologists face. RT research also deals with image production, patient interaction, radiation safety, quality assurance, positioning, protocol adherence, equipment operation, workflow coordination, clinical education, and the translation of policy into practice. In the AI era, these domains become methodological sites rather than background details.

Radiologic Technology is used here in a broad professional sense. It includes diagnostic radiography, computed tomography, magnetic resonance imaging, mammography, fluoroscopy, interventional imaging support, operating-theater and mobile imaging, radiation safety, image quality assurance, digital imaging systems, departmental workflow, and RT education. The term also covers the educational and research functions of clinical instructors, thesis advisers, program heads, academic leaders, and radiographers who conduct practice-based inquiry. The review therefore speaks to both clinical and academic researchers, especially in settings where AI adoption may vary by infrastructure, budget, governance, vendor support, and staff training.

The manuscript uses a narrative and conceptual review approach. A systematic review would be appropriate if the goal were to estimate pooled diagnostic accuracy or summarize a tightly defined intervention. The present goal is different. It synthesizes methodological concerns across reporting guidelines, diagnostic accuracy research, imaging AI validation, human factors, implementation science, education, ethics, and radiography workforce literature. This approach suits a field that needs conceptual organization before it can standardize research protocols. It also permits a framework-building contribution rather than a narrow summary of individual AI tools.

The synthesis is informed by the broader medical AI literature that has cautioned against treating algorithm development as equivalent to clinical impact. Beam and Kohane (2018) framed machine learning as a major shift in health-care data use, while Topol (2019) argued that high-performance medicine depends on a careful convergence of human and artificial intelligence. Similar arguments appear in radiology and medical imaging scholarship, where Pesapane et al. (2018), Thrall et al. (2018), and Rajpurkar et al. (2022) emphasized both opportunity and implementation risk. Regulatory and commercial evidence also remains uneven, as shown by Benjamens et al. (2020), Muehlematter et al. (2021), and van Leeuwen et al. (2021). For RT researchers, these sources support a practical conclusion: evidence should not stop at proof that an algorithm can run; it should show that the tool improves a relevant imaging process under the conditions in which staff actually work.

The review draws on peer-reviewed journal literature published from 2016 onward. This date range reflects the contemporary era of deep learning, clinical AI reporting standards, and serious discussion of real-world deployment. Foundational studies on deep learning in medical imaging, clinical trial reporting, bias, dataset shift, external validation, and radiography professional impact are used only where they directly support the methodological argument (Esteva et al., 2019; Hosny et al., 2018; Litjens et al., 2017; Park & Han, 2018). The emphasis remains on the current research problem: how to design RT studies that can survive peer review, inform practice, and remain useful after AI becomes part of routine imaging environments.

A conceptual boundary is necessary. The review does not teach machine learning programming, compare neural network architectures, or rank commercial AI products. It does not reduce RT research to radiologist outcome measures. Instead, it asks what should count as valid evidence when AI intersects with image acquisition, protocol choice, dose, positioning, image quality, radiographer behavior, workflow, education, and departmental accountability. This positioning makes the manuscript distinct from a generic AI-in-radiology review.

3. Current AI Reporting Standards and Their Relevance to Radiologic Technology Research

Recent reporting standards have improved the quality expectations for AI research in medicine. CONSORT-AI and SPIRIT-AI require investigators to describe AI interventions in clinical trial protocols and reports, including input data, human-AI interaction, error analysis, and integration into clinical pathways (Cruz Rivera et al., 2020; Liu et al., 2020). DECIDE-AI supports early-stage clinical evaluation before large randomized trials, a stage that is highly relevant for imaging departments that test AI in controlled or limited clinical settings (Vasey et al., 2022). TRIPOD+AI strengthens prediction-model reporting by requiring clearer information on data, intended use, model development, validation, and performance evaluation (Collins et al., 2024). STARD-AI extends diagnostic accuracy reporting for AI-enabled tests, including additional items on fairness, generalizability, and human-AI interaction (Sounderajah et al., 2025). CLAIM 2024 provides medical-imaging-specific guidance for AI manuscripts and reflects lessons from the rapid expansion of imaging AI research (Tejani et al., 2024). FUTURE-AI adds a broader trustworthy-AI lens, emphasizing fairness, universality, traceability, usability, robustness, and explainability (Lekadir et al., 2025).

These standards matter for Radiologic Technology research because they shift the expectation from impressive model performance to transparent, reproducible, clinically meaningful evidence. They ask authors to define the intended use, describe the data source, report validation procedures, explain the AI intervention, and consider human interaction. For an RT researcher, these requirements are immediately relevant. A study that evaluates AI-assisted chest radiography positioning, CT dose modulation, image quality alerts, or MRI protocol recommendation must not only report algorithmic output. It must describe how images were acquired, who interacted with the tool, what decisions changed, and whether patient care or workflow improved.

Reporting quality also depends on tools that predate the latest AI-specific extensions. Cohen et al. (2016) clarified diagnostic accuracy reporting through STARD 2015, while Moons et al. (2019) and Steyerberg and Harrell (2016) emphasized risk-of-bias assessment and validation in prediction modeling. Mongan et al. (2020) provided the original CLAIM checklist for imaging AI before the 2024 update, and Wynants et al. (2020) illustrated how weak reporting and validation can compromise urgent clinical prediction research. These sources matter because RT-AI studies can appear novel while still repeating old methodological errors: unclear reference standards, narrow development datasets, inadequate validation, and outcome claims that exceed the design.

At the same time, existing AI reporting standards are not designed specifically around radiologic technologist-centered variables. They do not always force authors to report exposure index, dose-area product, CTDIvol, DLP, positioning criteria, reject-analysis thresholds, portable imaging conditions, breath-hold instruction, protocol deviations, staffing conditions, patient transfer limitations, or radiographer override behavior. CLAIM 2024 and STARD-AI strengthen imaging and diagnostic accuracy reporting, but RT research still needs more explicit adaptation for acquisition conditions and human workflow (Sounderajah et al., 2025; Tejani et al., 2024). A model trained on images produced under one technical culture may not behave the same way in another department that uses different equipment, exposure habits, or patient preparation procedures.

The implication is not that RT researchers need separate standards that ignore existing guidance. Rather, they need an RT-aware extension of current standards. A thesis, institutional study, or journal article should use CONSORT-AI, SPIRIT-AI, DECIDE-AI, TRIPOD+AI, STARD-AI, CLAIM 2024, or FUTURE-AI when relevant, but it should add acquisition and workflow details that matter in radiography practice. The method section should answer practical questions: Who acquired the images? Under which protocol? With what equipment? Under what positioning criteria? How were repeats handled? Did AI feedback occur before, during, or after acquisition? Could radiographers override the system? Was training provided? Were local workflow consequences measured?

This adaptation would move RT-AI manuscripts closer to the evidence standard expected by reputable journals. It would also prevent a common weakness: a technically sound AI study with insufficient clinical and acquisition context. In Radiologic Technology, context is not a decorative paragraph. It is part of the validity structure of the study.

4. Methodological Gaps in AI-Enabled Radiologic Technology Research

The first methodological gap is the persistence of accuracy-centered evaluation. AUC, sensitivity, specificity, precision, recall, Dice score, and F1-score are important, but they do not establish whether an AI tool improves radiography practice. Several reviews of clinical AI have noted that many studies remain retrospective, internally validated, and insufficiently tested in real clinical environments (Liu et al., 2019; Nagendran et al., 2020). Radiologic Technology research must therefore ask questions that model metrics alone cannot answer. Does the tool reduce repeat imaging? Does it improve positioning consistency? Does it reduce dose while preserving diagnostic acceptability? Does it shorten examination time without compromising patient safety? Does it interrupt workflow or create additional verification steps?

The second gap is weak external validation. Imaging AI performance can deteriorate when models move across institutions, scanners, patient groups, disease prevalence, or acquisition protocols. Yu et al. (2022) emphasized that external validation remains infrequent in radiology AI studies and that performance often declines outside development settings. This problem has direct RT implications. A model trained on high-quality, fixed-room radiographs from a large academic hospital may not perform similarly on portable intensive-care-unit images, locally adjusted exposure protocols, older detectors, or patient populations with different body habitus and disease patterns.

The third gap is poor documentation of acquisition conditions. AI datasets often describe modality and diagnosis, but omit the technical details that radiographers would recognize as decisive. Positioning, centering, collimation, grid use, source-to-image distance, exposure index, motion, artifacts, protocol deviations, and equipment differences can influence both image quality and algorithm output. If these variables are not reported, readers cannot judge whether the findings can transfer to their own departments. This gap is especially serious for studies that claim implications for workflow or quality assurance.

The fourth gap concerns human factors. AI does not enter a department as a neutral calculator. It changes attention, trust, workload, decision confidence, and accountability. Automation bias, alert fatigue, cognitive offloading, and overreliance can arise when users treat AI outputs as more reliable than the situation warrants

Table 1. Major methodological gaps in AI-enabled Radiologic Technology research and proposed corrective strategies.

Methodological gap	Why it matters in RT research	Corrective strategy
Accuracy-centered evaluation	Model metrics do not prove value for dose, image quality, repeat rate, workflow, or patient care.	Add RT-relevant outcomes such as repeat rate, positioning adequacy, dose indices, workflow time, workload, and patient experience.
Weak external validation	AI performance may decline across equipment, protocols, patient groups, and institutions.	Use multicenter or external test sets; report equipment, protocol, and subgroup performance.
Limited acquisition documentation	Images are shaped by positioning, exposure, artifacts, patient cooperation, and local protocols.	Include acquisition-context tables with modality, protocol, exposure or dose indices, positioning criteria, equipment, and reject handling.
Insufficient human-factors evidence	Radiographers may overtrust, distrust, ignore, or misinterpret AI outputs.	Use simulation, usability testing, think-aloud protocols, time-motion analysis, and mixed-methods interviews.
Limited implementation evidence	AI must fit PACS/RIS, staffing, training, QA, and departmental workflow.	Measure acceptability, feasibility, adoption, fidelity, cost, sustainability, and integration burden.
Unclear accountability	Errors may involve AI output, user judgment, workflow design, or vendor limitations.	Report decision pathways, override rules, training, governance, audit procedures, and incident response.
Overreliance on perception surveys	Readiness studies do not show whether AI changes practice safely or effectively.	Move toward validation, implementation, education, human-AI interaction, and quality-assurance designs.

(Cabitza et al., 2017; Challen et al., 2019; Park, 2025). Conversely, distrust can lead users to ignore useful alerts. Radiographer-centered AI research must therefore study how staff interpret AI feedback, when they override it, how training affects trust calibration, and whether interface design supports safe decision-making.

The fifth gap is limited implementation evidence. Real departments must integrate AI with picture archiving and communication systems, radiology information systems, electronic health records, staffing patterns, reporting pathways, and quality assurance structures. Reviews of AI adoption stress the importance of interoperability, production workflows, feedback systems, and monitoring (Dikici et al., 2020; Tejani et al., 2024). Yet RT studies often stop at attitudes or simulated performance. Stronger work should evaluate feasibility, adoption, fidelity, sustainability, cost, and safety monitoring after implementation.

The sixth gap concerns local generalizability and equity. AI models can reproduce or amplify bias when training data do not reflect the populations and settings where the tool is used. Studies have shown that medical AI can encode demographic signals and that algorithmic performance may differ across groups or settings (Gichoya et al., 2022; Obermeyer et al., 2019; Ricci Lara et al., 2022; Seyyed-Kalantari et al., 2021). RT researchers in the Philippines, ASEAN, and other low- and middle-income contexts have a responsibility and an opportunity to test whether AI tools remain valid under local equipment, patient, staffing, and workflow conditions.

5. Image Acquisition as a Core Methodological Variable

Image acquisition should be treated as a core methodological variable in AI-enabled Radiologic Technology research. In conventional retrospective AI studies, images are often treated as finished data objects. Radiographers know that every image is the end point of a process. Patient identification, preparation, explanation, immobilization, positioning, exposure selection, protocol selection, detector placement, breathing instruction, and post-processing decisions all shape the image before it becomes data. If those steps vary across departments, the dataset varies in ways that can influence algorithmic performance.

Positioning is an obvious example. AI tools that assess image quality, detect anatomy, estimate pathology, or triage urgent findings may respond differently when anatomy is rotated, clipped, underinflated, overexposed, underexposed, or obscured by artifacts. A study that evaluates AI for chest radiography but does not report positioning criteria, inspiratory adequacy, portable status, projection, or repeat policy leaves a serious interpretive gap. In mammography, compression, positioning quality, tissue inclusion, and view adequacy can influence both diagnostic confidence and AI output. In CT, centering, patient size, scan range, reconstruction kernel, slice thickness, and dose modulation can alter image texture and model behavior. In MRI, coil selection, sequence parameters, patient motion, and protocol deviations can shape downstream segmentation or classification performance.

The technical imaging literature supports this acquisition-aware view. Shen et al. (2017), Litjens et al. (2017), and Lundervold and Lundervold (2019) showed that deep learning in medical image analysis depends heavily on image characteristics, preprocessing, and modality-specific structure. Hwang et al. (2019) demonstrated the feasibility of deep learning for chest radiograph abnormality detection, while McKinney et al. (2020) and Lång et al. (2023) showed that breast imaging AI can achieve clinically important performance under defined study conditions. These studies should not be read only as evidence of AI capability. For Radiologic Technology, they also show why acquisition, equipment, image quality, and screening workflow must be documented when authors claim clinical relevance.

Radiation dose variables also deserve stronger methodological status. AI can support dose optimization through reconstruction, protocol recommendation, exposure monitoring, or quality feedback, but dose-related claims require clear outcome definitions. In projection radiography, exposure index, deviation index, dose-area product, and repeat rate can be relevant. In CT, CTDIvol, DLP, size-specific dose estimates, scan length, and reconstruction method should be considered. A study that reports diagnostic accuracy while ignoring dose may not address the central safety concerns of RT practice. Conversely, a study that reports dose reduction without diagnostic acceptability may overstate benefit.

Equipment and vendor differences can create hidden domain shifts. Detector technology, image processing algorithms, reconstruction engines, software versions, calibration practices, display conditions, and local protocol naming conventions differ across institutions. AI tools trained under one technical configuration may not generalize to another. Willemink et al. (2020) stressed that preparing imaging data for machine learning requires careful attention to data heterogeneity, annotation, and preprocessing. For RT researchers, that attention must extend upstream to acquisition and quality control. The machine-learning dataset begins at the examination room, not at the spreadsheet.

Portable imaging provides a useful example of why acquisition context matters. Bedside radiography is often performed under constraints: limited patient mobility, medical devices, infection-control precautions, crowded spaces, suboptimal detector positioning, and urgent clinical requests. These conditions can increase variability and artifact burden. If AI tools are tested only on stable, well-positioned images, their apparent performance may not reflect the environment where they are most needed. RT research should therefore document fixed versus portable acquisition, emergency versus elective context, patient cooperation, and repeat-imaging policy when such variables may affect performance.

A practical recommendation follows: AI-enabled RT studies should include an acquisition-context table. This table should list modality, projection or protocol, equipment type, vendor or software version where appropriate, exposure or dose indices, positioning criteria, image quality criteria, repeat/reject handling, patient preparation requirements, acquisition setting, and staff training. Such a table would strengthen reproducibility and make RT contributions visible in AI research.

6. Beyond Accuracy: Radiologic Technology-Relevant Outcome Measures

AI evaluation in Radiologic Technology must move beyond diagnostic accuracy. Accuracy remains necessary when the tool detects, classifies, predicts, or prioritizes findings, but it is not sufficient. Radiologic technologists practice in a domain where the value of an intervention may lie in fewer repeats, better positioning, improved dose efficiency, faster examination flow, reduced patient movement, clearer image-quality decisions, safer escalation, or more consistent education. A model can be accurate in a test dataset and still be impractical, disruptive, unsafe, or poorly matched to departmental workflow.

Image quality outcomes should include diagnostic acceptability, artifact rate, noise, contrast, positioning adequacy, collimation, anatomy inclusion, rejected image proportion, and repeat reason. These outcomes connect AI evaluation to existing quality assurance practices. A tool that alerts radiographers to motion, rotation, clipped anatomy, or exposure deviation should be judged by whether it reduces clinically meaningful repeats, not only by whether it identifies technical errors retrospectively. Antun et al. (2020) warned that deep-learning methods can be unstable under small image perturbations, which reinforces the need to examine technical image variation in practice.

Radiation safety outcomes should be equally explicit. Depending on modality, studies may measure exposure index, deviation index, dose-area product, entrance surface dose estimates, CTDIvol, DLP, size-specific dose estimate, fluoroscopy time, cumulative air kerma, or dose reduction at maintained diagnostic quality. Chen et al. (2017) demonstrated deep-learning reconstruction for low-dose CT, and many subsequent studies have explored AI-enabled dose reduction or reconstruction. For RT research, the methodological issue is not only whether AI can produce a visually acceptable image at lower dose, but whether the study defines dose outcomes, image-quality thresholds, repeat consequences, and patient safety trade-offs with enough clarity.

Workflow outcomes deserve more attention. Examination time, patient throughput, reporting turnaround, urgent-case prioritization, interruption frequency, communication delay, image-transfer failures, and time spent resolving AI alerts can all shape departmental value. Tejani et al. (2024) argued that AI integration in radiology requires standards-based interoperability and workflow design. Dikici et al. (2020) similarly proposed maturity levels for research, production, and feedback in AI workflow integration. RT researchers can adapt these ideas by measuring whether AI changes radiographer tasks, handoffs, and quality checks rather than assuming that model performance translates to efficiency.

Human-factors outcomes are also critical. Trust, usability, cognitive workload, automation bias, alert fatigue, decision confidence, override rate, explanation usefulness, and perceived accountability should be measured when radiographers interact with AI. Chen et al. (2024) found that human-AI collaboration can reduce workload and improve efficiency in some medical contexts, but such effects depend on the task, interface, and clinical environment. RT studies should not assume that AI assistance is automatically helpful. A well-designed study may combine quantitative workload scales, time-motion data, error analysis, and qualitative interviews to show how users actually experience the tool.

A wider human-AI literature also warns that collaboration effects are task-specific. Tschandl et al. (2020) showed that human-computer collaboration can change diagnostic performance, but the direction and magnitude of benefit depend on expertise and presentation of AI advice. Wiens et al. (2019) framed responsible machine learning as a clinical safety problem, and Panch et al. (2019) cautioned that poor data and poor implementation can limit health-care AI. These lessons apply directly to radiography workflows, where the user may need to act quickly, communicate with the patient, and balance image ideality against clinical practicality.

Patient-centered outcomes should not disappear from AI methodology. Patient waiting time, communication quality, privacy concerns, need for repeat exposure, positioning discomfort, anxiety, and satisfaction may be relevant depending on the tool. An AI-assisted positioning system that improves image quality but prolongs setup time or reduces patient communication may create mixed effects. A triage tool that speeds urgent interpretation may also change how radiographers communicate uncertainty. RT research should therefore frame AI as a patient-care intervention as well as a technical innovation.

7. The Radiographer-in-the-Loop: Human-AI Interaction in Imaging Workflows

The radiographer-in-the-loop concept should become a central methodological lens in RT-AI research. Current human-AI discussions in radiology often focus on radiologist interpretation. That focus is understandable, but incomplete. Radiologic technologists interact with AI at different points: before acquisition through protocol recommendation or scheduling prioritization, during acquisition through positioning or exposure guidance, immediately after acquisition through image-quality alerts, and later through reject analysis, dose monitoring, and workflow escalation. Each point creates a different research question.

Human-AI interaction must be studied as behavior, not as attitude alone. Surveys can capture readiness and concern, and studies from Ghana, Saudi Arabia, and other contexts have shown that radiographers recognize both opportunities and professional risks in AI adoption (Aldhafeeri, 2022; Botwe et al., 2021). However, readiness scores do not show how radiographers respond to an AI alert when a patient is in pain, how they handle conflicting cues between experience and algorithmic recommendation, or how they decide whether an image is acceptable after AI feedback. Those questions require observation, simulation, usability testing, think-aloud protocols, eye tracking, time-motion analysis, or mixed-methods designs.

Automation bias is a key risk. A radiographer may accept AI feedback because it appears objective, even when the clinical situation suggests caution. The opposite risk is algorithm aversion, where users reject a tool after seeing errors or after receiving inadequate training. Both patterns can reduce safety. Challen et al. (2019) described potential sources of failure in machine-learning systems, including implementation and human oversight issues. Cabitza et al. (2017) also warned against uncritical use of machine learning in clinical decision support. RT research should therefore measure not only the correctness of AI outputs but also how human users respond to correct, incorrect, uncertain, and ambiguous outputs.

Alert fatigue deserves careful study. If an image-quality algorithm produces frequent low-value alerts, radiographers may ignore it. If a dose-monitoring tool flags many non-actionable events, staff may treat alerts as administrative noise. If a triage system escalates too many cases, it may disrupt workflow without improving care. RT-AI studies should record alert frequency, actionability, response time, override rate, false-positive burden, and user perception of alert usefulness. These outcomes can connect human factors to measurable workflow consequences.

Explainability should be evaluated in practical terms. A heatmap or probability score may satisfy a technical definition of explanation but still fail to help a radiographer decide whether to repeat an image, adjust exposure, reposition the patient, or proceed with the examination. Explanations should be tested against the decisions users actually make. A study could compare different feedback formats, such as a simple pass/fail quality alert, an annotated anatomical marker, a positioning score, or a ranked list of technical deficiencies. The outcome should be safe and consistent action, not only user preference.

The radiographer-in-the-loop model also respects professional judgment. It does not suggest that radiographers should defer to AI. It suggests that their interaction with AI is itself a research object. Experienced radiographers carry tacit knowledge about patient cooperation, trauma limitations, pediatric behavior, emergency constraints, and the trade-off between ideal images and acceptable images. AI research that ignores this knowledge risks designing tools that look impressive in a laboratory and fail in practice.

8. Dataset Quality, Bias, and Local Generalizability

Dataset quality is one of the strongest determinants of AI validity. In Radiologic Technology research, dataset quality includes more than file format, pixel resolution, or label availability. It includes the conditions under which images were acquired, the consistency of protocols, the quality of annotations, the representativeness of patient groups, and the presence of technical artifacts that reflect real practice. A model trained on a clean dataset may not remain reliable when applied to images with portable artifacts, local positioning variation, older equipment, or different processing pipelines.

Label validity requires careful attention. Imaging labels may come from radiology reports, expert review, consensus panels, pathology, clinical follow-up, laboratory results, or administrative codes. Each source has limitations. Reports may contain uncertainty. Clinical follow-up may be incomplete. Expert labels may vary. Inter-rater agreement should be reported when labels involve judgment, especially for image quality, positioning adequacy, artifact severity, and technical acceptability. CLAIM 2024 and STARD-AI both reinforce the need for transparent reporting of reference standards and evaluation processes in AI imaging research (Sounderajah et al., 2025; Tejani et al., 2024).

Class imbalance can distort results. A tool that detects uncommon positioning errors, rare complications, or low-frequency urgent findings may appear accurate because negative cases dominate the dataset. Metrics should therefore match the clinical problem. Precision, recall, calibration, decision-curve analysis, subgroup performance, and false-negative analysis may be more informative than headline accuracy. For RT researchers, error analysis should include technical and workflow consequences. A false reassurance from an AI quality tool may lead to a missed repeat opportunity; an unnecessary alert may delay patient throughput.

Domain shift is central to local generalizability. Differences in equipment, image processing, patient population, disease prevalence, staff training, acquisition protocol, and workflow can change the input distribution. Zech et al. (2018) showed that models trained on chest radiographs can learn institution-specific patterns that limit generalizability. DeGrave et al. (2021) later showed that shortcut learning can mislead medical imaging models. Kore et al. (2024) and Sahiner et al. (2023) emphasized that data drift and dataset shift can degrade deployed systems. RT researchers should therefore treat external validation as a requirement, not a luxury.

Bias and fairness are not abstract issues. Gichoya et al. (2022) showed that AI can recognize self-reported race from medical images, even when humans cannot easily identify the mechanism. Obermeyer et al. (2019) documented racial bias in a widely used health algorithm, and Seyyed-Kalantari et al. (2021) showed underdiagnosis patterns in chest radiography models. These studies caution that AI performance may differ across demographic groups and clinical contexts. In RT research, fairness analysis may require stratification by age, sex, body habitus, disability, equipment type, care setting, or institution type, depending on the question.

The need for local validation is especially strong in Philippine, ASEAN, and other resource-variable settings. AI tools developed in high-resource systems may assume infrastructure, image quality, data availability, staffing, and connectivity that are not universal. Local validation does not imply distrust of global innovation; it reflects methodological responsibility. RT researchers can contribute by building multicenter local datasets, testing AI under real acquisition conditions, documenting equipment heterogeneity, and reporting subgroup performance in a way that helps departments decide whether adoption is safe and worthwhile.

Health data poverty adds another layer to this problem. Ibrahim et al. (2021) described data poverty as a barrier to equitable digital health care, while Cross et al. (2024) treated bias in medical AI as a health-equity concern. Image data augmentation can help technical development, as reviewed by Shorten and Khoshgoftaar (2019), but augmentation cannot replace local evidence about acquisition conditions, patient characteristics, or workflow constraints. RT researchers should therefore distinguish between technical attempts to diversify data and clinical attempts to validate performance in the setting where decisions will occur.

9. Implementation Science and Post-Deployment Monitoring

AI adoption is often described as if the main challenge were model development. In practice, the harder challenge may be implementation. An algorithm must enter a department with existing PACS, RIS, EHR, modality worklists, quality assurance procedures, reporting pathways, staffing constraints, and professional cultures. Implementation science gives RT researchers a vocabulary for this problem. Instead of asking only whether an AI tool works, it asks whether the tool is acceptable, feasible, adopted, implemented with fidelity, sustained over time, and worth its cost.

Kelly et al. (2019) identified delivery of clinical impact as one of the central challenges for health-care AI. That warning is especially relevant to Radiologic Technology because the distance between algorithm output and patient benefit often passes through several practical steps: acquisition, image review, communication, documentation, handoff, and quality assurance. A model may perform well in a validation dataset but fail to produce value if radiographers do not receive the result in time, if alerts appear outside their usual workstation, if training is superficial, or if departmental policy does not define who should act on the recommendation. Implementation science converts these practical issues into researchable variables.

Implementation outcomes should be selected before adoption. Acceptability can measure whether radiographers, radiologists, administrators, and patients consider the tool appropriate. Feasibility can examine whether the tool fits available infrastructure and staffing. Adoption can measure actual use. Fidelity can test whether users apply the tool as intended. Cost can include software, hardware, training, maintenance, integration, staff time, and opportunity cost. Sustainability can assess whether the tool continues to provide value after the early enthusiasm ends. These outcomes are highly relevant to institutions that must justify AI investment.

Workflow integration is a technical and social problem. Tejani et al. (2024) emphasized standards-based interoperability for radiology AI, while Dikici et al. (2020) described research, production, and feedback maturity. In RT departments, integration must also account for who receives the alert, where the alert appears, whether it interrupts acquisition, whether it changes a protocol, how it is documented, and who is responsible for follow-up. A technically accurate alert that appears outside the radiographer’s normal workflow may be ignored. A quality tool that requires duplicate data entry may fail despite good performance.

Post-deployment monitoring is essential. AI systems can degrade because of scanner upgrades, protocol changes, population shifts, new disease patterns, vendor software updates, changes in staff practice, or data pipeline errors. Dean et al. (2025) stressed that deployed radiology AI requires monitoring across performance, reliability, workflow, and safety. Kore et al. (2024) showed that data drift can affect medical machine learning systems. RT research can contribute by defining what should be monitored after deployment: repeat rates, dose distributions, alert burden, false positives, false negatives, user overrides, workflow delays, subgroup performance, and safety incidents.

Governance should be part of methodology. A study that introduces AI into a department should identify oversight responsibility, escalation procedures, audit intervals, training requirements, user-feedback channels, documentation rules, and criteria for pausing or recalibrating a tool. The governance plan can be simple in a small study and more formal in a multicenter deployment, but it should not be absent. Without governance, AI becomes an unmanaged variable inside a safety-critical imaging environment.

RT researchers are well positioned to study implementation because they understand the operational realities of imaging departments. They know when a proposed workflow adds time, when a protocol recommendation is impractical, when a quality alert conflicts with patient condition, and when staff need training rather than more automation. This field knowledge should be treated as methodological strength.

10. Ethics, Equity, Accountability, and AI Literacy

Ethical AI in Radiologic Technology begins with the patient. Imaging data are personal health data, and AI development often relies on secondary use of archived images, metadata, reports, and clinical outcomes. Researchers must address consent, privacy, de-identification, data governance, cybersecurity, and institutional approval. Vayena et al. (2018) argued that machine learning in medicine requires careful attention to data ethics and trust. Char et al. (2018) similarly emphasized ethical challenges related to clinical machine learning. For RT researchers, the ethical question includes not only data use but also patient exposure, repeat imaging, communication, and the professional responsibility of staff who act on AI feedback.

Equity should be studied, not assumed. AI tools may perform differently across demographic groups, equipment types, body habitus, care settings, and institutions. Fairness analysis should therefore be built into study design when the sample size allows. If a study cannot test fairness adequately, it should state that limitation clearly. Claims of generalizability should be modest unless supported by external validation and subgroup analysis. This caution is essential in settings where imported AI tools may be trained on populations and equipment unlike local practice.

Accountability becomes complex when AI influences technical decisions. If an AI tool recommends a protocol, flags poor positioning, or triages a case, who is responsible when the recommendation is wrong or misunderstood? The answer may involve the radiographer, radiologist, department, vendor, institution, and regulator. RT research cannot resolve all legal questions, but it can document decision pathways and user behavior. Studies should record whether AI output was advisory or directive, whether users could override it, how overrides were logged, and what training users received.

Explainability is also an ethical issue. Users need enough information to judge when AI output should be trusted. However, explainability must be matched to the user and task. A radiographer may need a technical quality reason, an anatomical localization, a dose warning, or a protocol rationale, not a mathematical description of model architecture. RT studies can test which forms of explanation improve safe action and which create confusion or false confidence.

AI literacy must become part of RT education and research training. Radiographers do not need to become software engineers, but they need enough competence to appraise AI claims, understand validation, interpret performance metrics, identify bias, recognize workflow risks, and use AI responsibly. Recent education work in radiology has called for structured AI curricula and competency frameworks for trainees (Lindqwister et al., 2023; van Kooten et al., 2024). RT curricula should adapt these ideas to the work of image acquisition, radiation safety, quality assurance, patient care, and departmental workflow.

Radiography-specific scholarship has begun to make the same point. Hardy and Harvey (2020) mapped AI tasks against radiography work, while Malamateniou et al. (2024) argued that practice, education, and research must develop together as AI enters radiography. Shiang et al. (2022) showed that AI decision support can be incorporated into radiology training environments, and Geis et al. (2019) placed professional ethics at the center of radiology AI. For RT education, these sources suggest that AI literacy should not be a short add-on lecture. It should become part of research methods, clinical reasoning, image evaluation, quality assurance, and professional ethics.

Research writing itself now raises AI literacy questions. Generative AI can assist with language editing, literature organization, coding support, and brainstorming, but it can also introduce fabricated references, unsupported claims, and authorship ambiguity. RT researchers should disclose AI assistance when required by journals, verify every citation independently, and treat AI-generated text as draft material subject to scholarly judgment. This standard is particularly important for educators who supervise student research.

11. The RADIATE-AI Framework for Future Radiologic Technology Research

This review proposes the RADIATE-AI framework as a practical guide for designing, reviewing, and reporting AI-enabled Radiologic Technology studies. The framework is not a replacement for CONSORT-AI, SPIRIT-AI, DECIDE-AI, TRIPOD+AI, STARD-AI, CLAIM 2024, or FUTURE-AI. Rather, it translates their principles into the professional and methodological conditions of Radiologic Technology. It asks researchers to place image acquisition, human-AI interaction, outcome selection, implementation, ethics, education, and accountability at the center of study design.

The first component, Research question and real-world use case, asks whether the study addresses a genuine clinical, educational, or operational problem. AI should not be used merely because it is fashionable. A strong RT-AI study begins with a specific use case: reducing chest radiograph repeats, improving CT protocol selection, detecting positioning errors, monitoring dose indices, improving mammography quality, triaging urgent portable radiographs, supporting student learning, or strengthening departmental QA. The research question should specify the user, decision point, setting, and expected benefit.

The second component, Acquisition protocol and image-generation conditions, requires researchers to document how data were produced. This includes modality, projection or sequence, equipment, exposure or dose indices, positioning criteria, protocol deviations, image processing, portable status, and repeat/reject handling. This component makes visible the work that radiologic technologists perform before AI analysis begins. It also strengthens reproducibility and external validity.

The third component, Dataset quality and diversity, addresses provenance, labeling, representativeness, imbalance, missingness, and domain shift. A dataset should be described in enough detail that readers can judge whether it reflects the intended use. Labeling procedures should be transparent, and uncertainty should be acknowledged. Local diversity should include not only patient demographics but also equipment, acquisition setting, and protocol variation.

The fourth component, Interaction between radiographers and AI, treats user behavior as part of the intervention. Researchers should describe how AI output was displayed, when it appeared, whether it interrupted workflow, whether users could override it, and what training was provided. If the study involves human users, outcomes should capture trust, usability, workload, decision confidence, alert burden, and override behavior.

The fifth component, Appraisal beyond accuracy, requires outcome measures that match RT practice. Diagnostic accuracy may be relevant, but image quality, dose, repeat rate, positioning adequacy, workflow, safety, patient experience, education, and implementation outcomes may be equally important. This component prevents studies from overclaiming clinical value from model performance alone.

The sixth component, Translation and implementation, asks whether the tool can be integrated into real departments. It includes feasibility, acceptability, adoption, fidelity, sustainability, cost, interoperability, training, and stakeholder engagement. The seventh component, Ethics, equity, and education, requires attention to privacy, bias, accountability, disclosure, and AI literacy. The final component, Algorithmic accountability and impact monitoring, extends the study beyond launch. It requires audit procedures, drift monitoring, safety reporting, governance, and criteria for recalibration or discontinuation. Together, these components turn AI-related RT research from a technology topic into a mature methodological agenda.

Table 2. The RADIATE-AI framework for future Radiologic Technology research.

Component	Meaning	Methodological focus
R	Research question and real-world use case	Define the user, setting, decision point, intended benefit, and clinical or educational problem.
A	Acquisition protocol and image-generation conditions	Document modality, protocol, positioning, exposure or dose indices, equipment, artifacts, and repeat/reject handling.
D	Dataset quality and diversity	Report provenance, labels, reference standards, class balance, missing data, demographics, equipment, and domain shift.
I	Interaction between radiographers and AI	Describe AI display, training, user decisions, trust, workload, alert burden, and override behavior.
A	Appraisal beyond accuracy	Measure image quality, dose, workflow, safety, education, equity, implementation, and patient-centered outcomes.
T	Translation and implementation	Assess feasibility, acceptability, adoption, fidelity, cost, interoperability, and sustainability.
E	Ethics, equity, and education	Address privacy, bias, accountability, disclosure, AI literacy, and curriculum integration.
AI	Algorithmic accountability and impact monitoring	Plan audit, drift monitoring, safety reporting, user feedback, recalibration, and discontinuation criteria.

12. Recommended Study Designs and Outcome Measures

Different RT-AI questions require different designs. Retrospective diagnostic accuracy studies are appropriate when researchers want to test algorithmic performance on existing images, but they should include transparent dataset description, reference standards, subgroup analysis, and external validation when possible. Prospective observational studies suit questions about real-time AI use without randomization. They can examine how radiographers respond to alerts, whether repeats change, or how dose distributions shift after introduction of a tool.

Simulation-based studies are valuable when patient risk or workflow disruption must be controlled. A simulation can test AI-assisted positioning, image-quality decisions, protocol selection, or emergency escalation using standardized cases. It can measure decision time, correctness, confidence, workload, and response to erroneous AI outputs. Simulation also suits education research, especially when students or trainees need exposure to AI-assisted workflows before clinical use.

Human-AI interaction experiments can compare user performance with and without AI, different explanation formats, different alert thresholds, or different interface designs. These studies should avoid simplistic conclusions. A radiographer may improve on easy cases and worsen on ambiguous cases; a tool may reduce time but increase overreliance; an explanation may improve confidence without improving correctness. Outcomes should therefore include performance, workload, trust calibration, override behavior, and qualitative interpretation.

Mixed-methods implementation studies are suitable when the research question concerns adoption, feasibility, and sustainability. Quantitative data can measure use rates, time, repeat rate, dose, or alert burden. Qualitative interviews can explain why staff accept, adapt, ignore, or resist the tool. This design is especially useful in institutions where AI adoption intersects with budget limitations, staff workload, training gaps, or unstable infrastructure.

Delphi studies can help develop RT-specific AI competencies, reporting extensions, quality indicators, or curriculum outcomes. Expert consensus should not replace empirical validation, but it can define priorities for a field that lacks standardized guidance. Usability studies can evaluate interface design, navigation, cognitive workload, and error-prone interactions. Interrupted time-series designs can test whether AI introduction changes departmental trends in repeat rate, turnaround time, or dose indices while accounting for baseline patterns.

Cost-effectiveness and budget-impact studies remain uncommon but important. Department heads and administrators need evidence about total cost, training burden, maintenance, software licensing, hardware upgrades, time saved, repeats avoided, and downstream benefits. Multicenter validation studies are needed for generalizability, especially across equipment types and resource settings. Educational intervention studies can test AI literacy modules, simulation exercises, research-methods training, and critical appraisal activities. In short, RT-AI methodology should begin with the question and choose the design that can answer it, rather than forcing all AI topics into retrospective performance evaluation.

Table 3. Recommended study designs, suitable research questions, data sources, and RT-relevant outcome measures.

Study design	Suitable research question	Data sources	RT-relevant outcomes
Retrospective diagnostic accuracy study	Does the AI tool detect or classify a finding on existing images?	Archived images, reports, expert labels, clinical follow-up	Sensitivity, specificity, AUC, calibration, subgroup performance, false-negative analysis.
Prospective observational study	How does AI perform during routine imaging workflow?	Live cases, modality logs, PACS/RIS timestamps, user actions	Repeat rate, alert response, turnaround time, dose indices, protocol deviations.
Simulation-based study	How do radiographers use AI feedback under controlled cases?	Standardized cases, simulated workstations, structured scenarios	Decision accuracy, response time, trust, workload, override behavior.
Human-AI interaction experiment	Which interface or explanation format supports safer decisions?	Experimental cases, eye tracking, think-aloud data, questionnaires	Usability, trust calibration, cognitive load, error rate, action consistency.
Mixed-methods implementation study	Is AI feasible, acceptable, and sustainable in a department?	Use logs, staff surveys, interviews, focus groups, QA indicators	Adoption, fidelity, cost, barriers, facilitators, sustainability.
Interrupted time-series study	Did AI introduction change departmental trends?	Monthly QA reports, dose records, reject analysis, workflow data	Repeat rate trends, DLP/CTDIvol trends, throughput, alert burden.
Delphi study	What competencies or reporting items should RT-AI studies include?	Expert panels, iterative rating rounds, consensus feedback	Consensus indicators, priority domains, curriculum or reporting items.
Educational intervention study	Does AI literacy training improve student or staff competence?	Pretest/posttest scores, simulation performance, reflective outputs	AI literacy, critical appraisal, ethical reasoning, safe workflow decisions.

13. Future Research Agenda

A future research agenda for Radiologic Technology in the AI era should begin with an RT-specific reporting extension. This extension could adapt CLAIM 2024, STARD-AI, TRIPOD+AI, DECIDE-AI, and FUTURE-AI by adding acquisition variables, dose indicators, image-quality criteria, repeat/reject handling, radiographer interaction, training, and workflow outcomes. Such an extension would help thesis advisers, journal reviewers, and early-career researchers identify what must be reported in AI-enabled RT studies.

Table 4. Examples of outcome measures for AI-enabled Radiologic Technology studies.

Outcome domain	Examples of measures
Image quality	Diagnostic acceptability, artifact rate, positioning adequacy, anatomy inclusion, contrast, noise, reject reason.
Radiation safety	Exposure index, deviation index, dose-area product, CTDIvol, DLP, fluoroscopy time, repeat exposure.
Workflow	Examination time, turnaround time, patient throughput, interruption frequency, image transfer delay.
Human factors	Trust, usability, cognitive workload, alert fatigue, decision confidence, override rate.
Implementation	Acceptability, feasibility, adoption, fidelity, cost, sustainability, training completion.
Education	AI literacy, critical appraisal skill, simulation performance, ethical reasoning, research-method competence.
Equity	Performance by age, sex, body habitus, institution type, equipment type, portable versus fixed imaging.
Safety	Near-miss events, false-negative consequences, false-positive burden, incident reports, drift signals.

Local and multicenter datasets should be a second priority. The Philippines, ASEAN, and other resource-variable settings need imaging datasets that reflect local equipment, patient characteristics, disease patterns, protocol practices, and staffing realities. These datasets should not be built only for algorithm development. They should also support validation, fairness analysis, dose research, quality assurance, and education. Multicenter collaboration can reduce the risk that one institution's technical culture becomes mistaken for general evidence.

Radiographer-AI collaboration requires deeper study. Future research should examine how radiographers use AI during acquisition, how they interpret image-quality alerts, how they handle conflicting recommendations, and how expertise changes AI use. Studies should include novices, experienced radiographers, clinical instructors, and modality specialists. They should also examine how AI changes professional identity, supervision, and responsibility without reducing the analysis to anxiety about replacement.

Dose and repeat-imaging outcomes should become signature RT-AI research areas. AI tools that claim to optimize dose, reduce repeats, improve positioning, or assist quality control should be tested with RT-relevant endpoints. Researchers should measure not only technical image acceptability but also repeat reasons, dose consequences, patient experience, and workflow effects. Such studies would align AI research with long-standing radiography commitments to radiation protection and image quality.

AI literacy standards should be developed for RT education. Curricula should teach basic AI concepts, validation, bias, performance metrics, dataset quality, human-AI interaction, ethics, disclosure, and critical appraisal. Capstone projects and theses can move beyond awareness surveys toward small validation studies, simulation studies, usability tests, educational interventions, or implementation assessments. Faculty development is also needed, because educators cannot guide students in AI research if they lack methodological confidence.

Finally, governance and post-deployment monitoring should become routine research topics. Departments need evidence on how to audit AI performance, manage drift, document overrides, respond to incidents, retrain users, and decide whether a tool remains safe. RT researchers can lead studies that connect quality assurance, risk management, education, and clinical workflow. This agenda would make Radiologic Technology a contributor to trustworthy AI, not merely a recipient of tools developed elsewhere.

14. Conclusion

The age of artificial intelligence requires Radiologic Technology research to become more methodologically ambitious. The field should not limit its questions to whether AI is accurate or whether radiographers feel ready for change. Those questions remain useful, but they are no longer enough. RT research must ask whether AI is safe, usable, equitable, explainable, locally valid, educationally integrated, and beneficial to patients, radiographers, radiologists, and imaging departments.

This review has argued that radiologic technologists occupy a central position in AI-enabled imaging because they shape the image before it becomes data, the workflow before it becomes output, and the human response before it becomes clinical action. Their work influences acquisition quality, positioning, dose, patient preparation, image rejection, quality assurance, communication, and departmental safety. These realities should appear in research methodology, not remain hidden in the background.

The proposed RADIATE-AI framework offers one way forward. By linking real-world use cases, acquisition conditions, dataset quality, human-AI interaction, outcomes beyond accuracy, implementation, ethics, education, and accountability, the framework can guide RT theses, journal manuscripts, protocol development, institutional adoption studies, and curriculum reform. It also gives the field a language for evaluating AI as a socio-technical intervention rather than a stand-alone algorithm.

For experienced educators, clinical instructors, department heads, and researchers, the challenge is clear. Radiologic Technology must help define the evidence standards for AI in imaging practice. If the field contributes only perception studies, its methodological voice will remain limited. If it contributes acquisition-aware validation, human-centered evaluation, implementation evidence, and locally grounded accountability research, it can shape the future of trustworthy imaging AI.

Acknowledgment

Sincere appreciation is given to all peer reviewers for their valuable comments and suggestions, which helped the author to improve the quality of the manuscript.

Conflict of Interest Statement

The author declares no conflict of interest.

References

Aldhafeeri, F. M. (2022). Perspectives of radiographers on the emergence of artificial intelligence in diagnostic imaging in Saudi Arabia. *Insights into Imaging*, 13(1), 178. https://doi.org/10.1186/s13244-022-01319-z

Antun, V., Renna, F., Poon, C., Adcock, B., & Hansen, A. C. (2020). On instabilities of deep learning in image reconstruction and the potential costs of AI. *Proceedings of the National Academy of Sciences*, 117(48), 30088-30095. https://doi.org/10.1073/pnas.1907377117

Beam, A. L., & Kohane, I. S. (2018). Big data and machine learning in health care. *JAMA*, 319(13), 1317-1318. https://doi.org/10.1001/jama.2017.18391

Benjamens, S., Dhunnoo, P., & Mesko, B. (2020). The state of artificial intelligence-based FDA-approved medical devices and algorithms: An online database. *npj Digital Medicine*, 3, 118. https://doi.org/10.1038/s41746-020-00324-0

Botwe, B. O., Antwi, W. K., Arkoh, S., & Akudjedu, T. N. (2021). Radiographers' perspectives on the emerging integration of artificial intelligence into diagnostic imaging: The Ghana study. *Journal of Medical Radiation Sciences*, 68(3), 260-268. https://doi.org/10.1002/jmrs.460

Cabitza, F., Rasoini, R., & Gensini, G. F. (2017). Unintended consequences of machine learning in medicine. *JAMA*, 318(6), 517-518. https://doi.org/10.1001/jama.2017.7797

Challen, R., Denny, J., Pitt, M., Gompels, L., Edwards, T., & Tsaneva-Atanasova, K. (2019). Artificial intelligence, bias and clinical safety. *BMJ Quality & Safety*, 28(3), 231-237. https://doi.org/10.1136/bmjqs-2018-008370

Char, D. S., Shah, N. H., & Magnus, D. (2018). Implementing machine learning in health care - Addressing ethical challenges. *The New England Journal of Medicine*, 378(11), 981-983. https://doi.org/10.1056/NEJMp1714229

Chen, H., Zhang, Y., Kalra, M. K., Lin, F., Chen, Y., Liao, P., Zhou, J., & Wang, G. (2017). Low-dose CT with a residual encoder-decoder convolutional neural network. *IEEE Transactions on Medical Imaging*, 36(12), 2524-2535. https://doi.org/10.1109/TMI.2017.2715284

Chen, M., Wang, Y., Wang, Q., Shi, J., Wang, H., Ye, Z., Xue, P., & Qiao, Y. (2024). Impact of human and artificial intelligence collaboration on workload reduction in medical image interpretation. *npj Digital Medicine*, 7, 349. https://doi.org/10.1038/s41746-024-01328-w

Cohen, J. F., Korevaar, D. A., Altman, D. G., Bruns, D. E., Gatsonis, C. A., Hooft, L., Irwig, L., Levine, D., Reitsma, J. B., de Vet, H. C. W., & Bossuyt, P. M. M. (2016). STARD 2015 guidelines for reporting diagnostic accuracy studies: Explanation and elaboration. *BMJ Open*, 6(11), e012799. https://doi.org/10.1136/bmjopen-2016-012799

Collins, G. S., Moons, K. G. M., Dhiman, P., Riley, R. D., Beam, A. L., Van Calster, B., Ghassemi, M., Liu, X., Reitsma, J. B., van Smeden, M., & others (2024). TRIPOD+AI statement: Updated guidance for reporting clinical prediction models that use regression or machine learning methods. *BMJ*, 385, e078378. https://doi.org/10.1136/bmj-2023-078378

Cross, J. L., Mansfield, K. E., Platt, R. W., & others (2024). Bias in medical artificial intelligence: Implications for health equity. *PLOS Digital Health*, 3(11), e0000651. https://doi.org/10.1371/journal.pdig.0000651

Cruz Rivera, S., Liu, X., Chan, A. W., Denniston, A. K., Calvert, M. J., & SPIRIT-AI and CONSORT-AI Working Group (2020). Guidelines for clinical trial protocols for interventions involving artificial intelligence: The SPIRIT-AI extension. *Nature Medicine*, 26(9), 1351-1363. https://doi.org/10.1038/s41591-020-1037-7

Dean, J. C., Gichoya, J. W., Wiggins, W. F., & Safdar, N. M. (2025). Real-world monitoring of artificial intelligence in radiology. *Korean Journal of Radiology*, 26(1), 1-12. https://doi.org/10.3348/kjr.2025.0962

DeGrave, A. J., Janizek, J. D., & Lee, S. I. (2021). AI for radiographic COVID-19 detection selects shortcuts over signal. *Nature Machine Intelligence*, 3(7), 610-619. https://doi.org/10.1038/s42256-021-00338-7

Dikici, E., Bigelow, M., Prevedello, L. M., White, R. D., & Erdal, B. S. (2020). Integrating AI into radiology workflow: Levels of research, production, and feedback maturity. *Journal of Medical Imaging*, 7(1), 016502. https://doi.org/10.1117/1.JMI.7.1.016502

Esteva, A., Robicquet, A., Ramsundar, B., Kuleshov, V., DePristo, M., Chou, K., Cui, C., Corrado, G., Thrun, S., & Dean, J. (2019). A guide to deep learning in healthcare. *Nature Medicine*, 25(1), 24-29. https://doi.org/10.1038/s41591-018-0316-z

Geis, J. R., Brady, A. P., Wu, C. C., Spencer, J., Ranschaert, E., Jaremko, J. L., Langer, S. G., Borondy Kitts, A., Birch, J., Shields, W. F., van den Hoven van Genderen, R., Kotter, E., Wawira Gichoya, J., Cook, T. S., Morgan, M. B., Tang, A., Safdar, N. M., & Kohli, M. (2019). Ethics of artificial intelligence in radiology: Summary of the joint European and North American multisociety statement. *Radiology*, 293(2), 436-440. https://doi.org/10.1148/radiol.2019191586

Gichoya, J. W., Banerjee, I., Bhimireddy, A. R., Burns, J. L., Celi, L. A., Chen, L. C., Correa, R., Dullerud, N., Ghassemi, M., Huang, S. C., Kuo, P. C., Lungren, M. P., Palmer, L. J., Price, B. J., Purkayastha, S., Pyrros, A. T., Oakden-Rayner, L., Okechukwu, C., Seyyed-Kalantari, L., & Trivedi, H. (2022). AI recognition of patient race in medical imaging: A modelling study. *The Lancet Digital Health*, 4(6), e406-e414. https://doi.org/10.1016/S2589-7500(22)00063-2

Hardy, M., & Harvey, H. (2020). Artificial intelligence in diagnostic imaging: Impact on the radiography profession. *The British Journal of Radiology*, 93(1108), 20190840. https://doi.org/10.1259/bjr.20190840

Hosny, A., Parmar, C., Quackenbush, J., Schwartz, L. H., & Aerts, H. J. W. L. (2018). Artificial intelligence in radiology. *Nature Reviews Cancer*, 18(8), 500-510. https://doi.org/10.1038/s41568-018-0016-5

Hwang, E. J., Park, S., Jin, K. N., Kim, J. I., Choi, S. Y., Lee, J. H., Goo, J. M., Aum, J., Yim, J. J., Park, C. M., & others (2019). Development and validation of a deep learning-based automated detection algorithm for major thoracic diseases on chest radiographs. *JAMA Network Open*, 2(3), e191095. https://doi.org/10.1001/jamanetworkopen.2019.1095

Ibrahim, H., Liu, X., Zariffa, N., Morris, A. D., & Denniston, A. K. (2021). Health data poverty: An assailable barrier to equitable digital health care. *The Lancet Digital Health*, 3(4), e260-e265. https://doi.org/10.1016/S2589-7500(20)30317-4

Kelly, C. J., Karthikesalingam, A., Suleyman, M., Corrado, G., & King, D. (2019). Key challenges for delivering clinical impact with artificial intelligence. *BMC Medicine*, 17, 195. https://doi.org/10.1186/s12916-019-1426-2

Kore, A., Fleming, S. L., Moore, J. H., & others (2024). Data drift in medical machine learning: Implications and mitigation strategies. *Nature Communications*, 15, 1887. https://doi.org/10.1038/s41467-024-46142-w

Lång, K., Josefsson, V., Larsson, A. M., Larsson, S., Högberg, C., Sartor, H., Hofvind, S., Andersson, I., & Dustler, M. (2023). Artificial intelligence-supported screen reading versus standard double reading in the Mammography Screening with Artificial Intelligence trial. *The Lancet Oncology*, 24(8), 936-944. https://doi.org/10.1016/S1470-2045(23)00298-X

Lekadir, K., Frangi, A. F., Porras, A. R., Glocker, B., Cintas, C., Langlotz, C. P., & FUTURE-AI Consortium (2025). FUTURE-AI: Guiding principles and consensus recommendations for trustworthy artificial intelligence in health care. *BMJ*, 388, e081554. https://doi.org/10.1136/bmj-2024-081554

Lindqwister, A. L., Hassanpour, S., Lewis, P. J., & others (2023). AI-RADS: An artificial intelligence curriculum for residents. *Frontiers in Medical Technology*, 4, 1007708. https://doi.org/10.3389/fmedt.2022.1007708

Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., van der Laak, J. A. W. M., van Ginneken, B., & Sanchez, C. I. (2017). A survey on deep learning in medical image analysis. *Medical Image Analysis*, 42, 60-88. https://doi.org/10.1016/j.media.2017.07.005

Liu, X., Faes, L., Kale, A. U., Wagner, S. K., Fu, D. J., Bruynseels, A., Mahendiran, T., Moraes, G., Shamdas, M., Kern, C., Ledsam, J. R., Schmid, M. K., Balaskas, K., Topol, E. J., Bachmann, L. M., Keane, P. A., & Denniston, A. K. (2019). A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: A systematic review and meta-analysis. *The Lancet Digital Health*, 1(6), e271-e297. https://doi.org/10.1016/S2589-7500(19)30123-2

Liu, X., Cruz Rivera, S., Moher, D., Calvert, M. J., Denniston, A. K., & SPIRIT-AI and CONSORT-AI Working Group (2020). Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: The CONSORT-AI extension. *Nature Medicine*, 26(9), 1364-1374. https://doi.org/10.1038/s41591-020-1034-x

Lundervold, A. S., & Lundervold, A. (2019). An overview of deep learning in medical imaging focusing on MRI. *Zeitschrift fur Medizinische Physik*, 29(2), 102-127. https://doi.org/10.1016/j.zemedi.2018.11.002

Malamateniou, C., O'Regan, T., McFadden, S., & Jackson, S. (2024). Artificial intelligence in radiography: Practice, research and education. *Radiography*, 30(Suppl. 2), S56-S59. https://doi.org/10.1016/j.radi.2024.09.062

McKinney, S. M., Sieniek, M., Godbole, V., Godwin, J., Antropova, N., Ashrafian, H., Back, T., Chesus, M., Corrado, G. S., Darzi, A., Etemadi, M., Garcia-Vicente, F., Gilbert, F. J., Halling-Brown, M., Hassabis, D., Jansen, S., Karthikesalingam, A., Kelly, C. J., King, D., & Shetty, S. (2020). International evaluation of an AI system for breast cancer screening. *Nature*, 577(7788), 89-94. https://doi.org/10.1038/s41586-019-1799-6

Moons, K. G. M., Wolff, R. F., Riley, R. D., Whiting, P. F., Westwood, M., Collins, G. S., Reitsma, J. B., Kleijnen, J., & Mallett, S. (2019). PROBAST: A tool to assess risk of bias and applicability of prediction model studies. *Annals of Internal Medicine*, 170(1), 51-58. https://doi.org/10.7326/M18-1376

Mongan, J., Moy, L., & Kahn, C. E. (2020). Checklist for artificial intelligence in medical imaging (CLAIM): A guide for authors and reviewers. *Radiology: Artificial Intelligence*, 2(2), e200029. https://doi.org/10.1148/ryai.2020200029

Muehlematter, U. J., Daniore, P., & Vokinger, K. N. (2021). Approval of artificial intelligence and machine learning-based medical devices in the USA and Europe. *The Lancet Digital Health*, 3(3), e195-e203. https://doi.org/10.1016/S2589-7500(20)30292-2

Nagendran, M., Chen, Y., Lovejoy, C. A., Gordon, A. C., Komorowski, M., Harvey, H., Topol, E. J., Ioannidis, J. P. A., Collins, G. S., & Maruthappu, M. (2020). Artificial intelligence versus clinicians: Systematic review of design, reporting standards, and claims of deep learning studies. *BMJ*, 368, m689. https://doi.org/10.1136/bmj.m689

Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, 366(6464), 447-453. https://doi.org/10.1126/science.aax2342

Panch, T., Mattie, H., & Celi, L. A. (2019). The "inconvenient truth" about AI in healthcare. *npj Digital Medicine*, 2, 77. https://doi.org/10.1038/s41746-019-0155-4

Park, S. H. (2025). Human-AI interaction in radiology: Balancing automation, trust, and clinical judgment. *Korean Journal of Radiology*, 26(2), 120-126. https://doi.org/10.3348/kjr.2025.0071

Park, S. H., & Han, K. (2018). Methodologic guide for evaluating clinical performance and effect of artificial intelligence technology for medical diagnosis and prediction. *Radiology*, 286(3), 800-809. https://doi.org/10.1148/radiol.2017171920

Pesapane, F., Codari, M., & Sardanelli, F. (2018). Artificial intelligence in medical imaging: Threat or opportunity? Radiologists again at the forefront of innovation in medicine. *Insights into Imaging*, 9(5), 745-753. https://doi.org/10.1007/s13244-018-0647-0

Rajpurkar, P., Chen, E., Banerjee, O., & Topol, E. J. (2022). AI in health and medicine. *Nature Medicine*, 28(1), 31-38. https://doi.org/10.1038/s41591-021-01614-0

Ricci Lara, M. A., Echeveste, R., & Ferrante, E. (2022). Addressing fairness in artificial intelligence for medical imaging. *Nature Communications*, 13, 4581. https://doi.org/10.1038/s41467-022-32186-3

Sahiner, B., Chen, W., Pezeshk, A., & Petrick, N. (2023). Data drift in medical machine learning: Implications and potential remedies. *The British Journal of Radiology*, 96(1150), 20220878. https://doi.org/10.1259/bjr.20220878

Seyyed-Kalantari, L., Zhang, H., McDermott, M. B. A., Chen, I. Y., & Ghassemi, M. (2021). Underdiagnosis bias of artificial intelligence algorithms applied to chest radiographs in under-served patient populations. *Nature Medicine*, 27(12), 2176-2182. https://doi.org/10.1038/s41591-021-01595-0

Shen, D., Wu, G., & Suk, H. I. (2017). Deep learning in medical image analysis. *Annual Review of Biomedical Engineering*, 19, 221-248. https://doi.org/10.1146/annurev-bioeng-071516-044442

Shiang, T., Sarwar, A., Halabi, S. S., & others (2022). Implementation of an artificial intelligence decision support system in radiology residency training. *Clinical Imaging*, 92, 32-37. https://doi.org/10.1016/j.clinimag.2022.09.003

Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. *Journal of Big Data*, 6, 60. https://doi.org/10.1186/s40537-019-0197-0

Sounderajah, V., Ashrafian, H., Rose, S., Golub, R. M., Shetty, S., De Fauw, J., Karthikesalingam, A., Liu, X., Denniston, A. K., & others (2025). STARD-AI: An extension of the STARD 2015 reporting guideline for diagnostic accuracy studies involving artificial intelligence. *Nature Medicine*, 31, 1-8. https://doi.org/10.1038/s41591-025-03953-8

Steyerberg, E. W., & Harrell, F. E. (2016). Prediction models need appropriate internal, internal-external, and external validation. *Journal of Clinical Epidemiology*, 69, 245-247. https://doi.org/10.1016/j.jclinepi.2015.04.005

Tejani, A. S., Moy, L., Kahn, C. E., & others (2024). Checklist for artificial intelligence in medical imaging (CLAIM): 2024 update. *Radiology: Artificial Intelligence*, 6(4), e240300. https://doi.org/10.1148/ryai.240300

Tejani, A. S., Yi, P. H., Marin, D., & others (2024). Integrating and adopting artificial intelligence in the radiology workflow: A primer for standards and Integrating the Healthcare Enterprise profiles. *Radiology*, 311(3), e232653. https://doi.org/10.1148/radiol.232653

Thrall, J. H., Li, X., Li, Q., Cruz, C., Do, S., Dreyer, K., & Brink, J. (2018). Artificial intelligence and machine learning in radiology: Opportunities, challenges, pitfalls, and criteria for success. *Journal of the American College of Radiology*, 15(3), 504-508. https://doi.org/10.1016/j.jacr.2017.12.026

Topol, E. J. (2019). High-performance medicine: The convergence of human and artificial intelligence. *Nature Medicine*, 25(1), 44-56. https://doi.org/10.1038/s41591-018-0300-7

Tschandl, P., Rinner, C., Apalla, Z., Argenziano, G., Codella, N., Halpern, A., Janda, M., Lallas, A., Longo, C., Malvehy, J., Paoli, J., Puig, S., Rosendahl, C., Soyer, H. P., Sinz, C., & Kittler, H. (2020). Human-computer collaboration for skin cancer recognition. *Nature Medicine*, 26(8), 1229-1234. https://doi.org/10.1038/s41591-020-0942-0

van Kooten, M. J., de Vos, B. D., van Leeuwen, K. G., van Ginneken, B., & others (2024). A framework to integrate artificial intelligence training into radiology residency programs. *Insights into Imaging*, 15, 15. https://doi.org/10.1186/s13244-023-01595-3

van Leeuwen, K. G., Schalekamp, S., Rutten, M. J. C. M., van Ginneken, B., & de Rooij, M. (2021). Artificial intelligence in radiology: 100 commercially available products and their scientific evidence. *European Radiology*, 31(6), 3797-3804. https://doi.org/10.1007/s00330-021-07892-z

Vasey, B., Nagendran, M., Campbell, B., Clifton, D. A., Collins, G. S., Denaxas, S., Denniston, A. K., Faes, L., Geerts, B. F., Ibrahim, M., Liu, X., Mateen, B. A., Mathur, P., McCradden, M. D., Morgan, L., Ordish, J., Rogers, C., Saria, S., Ting, D. S. W., & McCulloch, P. (2022). Reporting guideline for the early-stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI. *Nature Medicine*, 28(5), 924-933. https://doi.org/10.1038/s41591-022-01772-9

Vayena, E., Blasimme, A., & Cohen, I. G. (2018). Machine learning in medicine: Addressing ethical challenges. *PLOS Medicine*, 15(11), e1002689. https://doi.org/10.1371/journal.pmed.1002689

Wiens, J., Saria, S., Sendak, M., Ghassemi, M., Liu, V. X., Doshi-Velez, F., Jung, K., Heller, K., Kale, D., Saeed, M., Ossorio, P. N., Thadaney-Israni, S., & Goldenberg, A. (2019). Do no harm: A roadmap for responsible machine learning for health care. *Nature Medicine*, 25(9), 1337-1340. https://doi.org/10.1038/s41591-019-0548-6

Willemink, M. J., Koszek, W. A., Hardell, C., Wu, J., Fleischmann, D., Harvey, H., Folio, L. R., Summers, R. M., Rubin, D. L., & Lungren, M. P. (2020). Preparing medical imaging data for machine learning. *Radiology*, 295(1), 4-15. https://doi.org/10.1148/radiol.2020192224

Wynants, L., Van Calster, B., Collins, G. S., Riley, R. D., Heinze, G., Schuit, E., Bonten, M. M. J., Damen, J. A. A., Debray, T. P. A., De Vos, M., Dhiman, P., Haller, M. C., Harhay, M. O., Henckaerts, L., Heus, P., Kreuzberger, N., Lohmann, A., Luijken, K., Ma, J., & van Smeden, M. (2020). Prediction models for diagnosis and prognosis of covid-19: Systematic review and critical appraisal. *BMJ*, 369, m1328. https://doi.org/10.1136/bmj.m1328

Yu, A. C., Eng, J., & Poon, A. D. (2022). External validation of deep learning algorithms for radiologic diagnosis: A systematic review. *Radiology: Artificial Intelligence*, 4(3), e210064. https://doi.org/10.1148/ryai.210064

Zech, J. R., Badgeley, M. A., Liu, M., Costa, A. B., Titano, J. J., & Oermann, E. K. (2018). Variable generalization performance of a deep learning model to detect pneumonia in chest radiographs: A cross-sectional study. *PLOS Medicine*, 15(11), e1002683. https://doi.org/10.1371/journal.pmed.1002683

Author Contributions: Alipio, M.; Study design, method conception, data collection, and manuscript writing.