LINK MEDICAL JOURNAL
of Health and Community Research | LMJHCR
Volume 3, Issue 2 | December 2025 | ISSN 3007-3448

DOI: 10.61919/gk9qxd94

Original Article

Clinician Trust in AI-Generated Radiology Narratives for Patient Care Decisions

Huang Yuzhe[1], Gilbert Lister[1], Armelia Adel Abdullah[1]

[1] Universitas Prima Indonesia

***Corresponding author: Gilbert Lister, gilbertlister@unprimdn.ac.id**

"Cite this Article" | **Received**: 23 August 2025; **Accepted**: 29 November 2025; **Published**: 30 December 2025.

ABSTRACT

Background: AI-generated radiology narratives are increasingly being considered for drafting, summarising, simplifying and organising radiology report information, but their safe use depends on how clinicians judge trust, uncertainty, accountability and workflow fit when patient-care decisions are involved. **Objective:** This qualitative study explored how clinicians judge the usefulness, limitations and safeguards of AI-generated radiology narratives in clinical decision-making. **Methods:** A qualitative design was used based on semi-structured interviews with ten clinicians from radiology, emergency medicine, oncology, respiratory medicine, orthopaedics, neurology, intensive care, cardiology and general practice. Participants discussed their experiences with radiology reports and their perceptions of AI-generated summaries, impressions, patient-friendly explanations and workflow-integrated narrative outputs. Reflexive thematic analysis was used to identify patterned meanings related to trust, risk, traceability, oversight and governance. **Results:** Five themes were generated: narrative usefulness as cognitive scaffolding rather than replacement; traceability and validation as the foundation of trust; accountability anxiety in higher-risk decisions; workflow fit and interprofessional communication; and provisional trust shaped by governance and learning. Clinicians valued AI-generated narratives when they organised lengthy reports, highlighted key findings and reduced cognitive load, but trust depended on clear labelling, source linkage, uncertainty preservation, radiologist verification, feedback routes, audit trails and local validation. Higher-risk decisions such as cancer staging, emergency discharge, surgical referral and critical-care escalation required stronger human oversight. **Conclusion:** Clinician trust in AI-generated radiology narratives is conditional, provisional and socio-technical. Safe implementation should prioritise transparent, traceable and governed narrative functions that support calibrated reliance rather than replacing radiologist judgement or professional accountability. **Keywords:** artificial intelligence; radiology narratives; clinician trust; qualitative research; thematic analysis; clinical decision-making; large language models; patient care; clinical governance.

EDITORIAL INFORMATION

Author Contributions: Concept: HY; Literature Review: GL; Drafting: AAA; Critical Revision and Final Approval: HY, GL, AAA.
Ethical Approval: Universitas Prima Indonesia
Informed Consent: Written informed consent was obtained from all participants
Conflict of Interest: The authors declare no conflict of interest; **Funding**: No external funding; **Data Availability**: Available from the corresponding author on reasonable request; **Acknowledgments**: N/A.

INTRODUCTION

Radiology has become one of the most information-intensive areas of contemporary healthcare, with computed tomography, magnetic resonance imaging, ultrasound, plain radiography and interventional imaging shaping diagnostic, therapeutic and referral decisions across emergency medicine, oncology, surgery, respiratory medicine, neurology, intensive care and primary care. This expanding dependence on imaging has intensified pressure on radiology services to deliver timely interpretation while also ensuring that imaging findings are communicated in a form that is clinically actionable for downstream care. In the

United Kingdom, the Royal College of Radiologists reported continuing workforce strain in clinical radiology, with its 2024 census noting that 344 doctors entered training and 265 completed training during that year (1). At the same time, artificial intelligence has moved from experimental development into increasingly regulated clinical products, with Singh et al. reviewing 1016 FDA-authorised AI-enabled medical devices and showing the growing scale of clinical AI adoption (2).

The challenge for radiology AI is not limited to whether an algorithm can detect or classify an abnormality. A radiology report is a professional clinical communication that can support diagnosis, justify discharge, trigger referral, guide treatment, recommend surveillance, influence multidisciplinary decision-making and shape discussion with patients. Recent radiology guidance therefore frames AI as part of the reporting environment rather than as an isolated technical add-on (3). Implementation evidence also shows that AI changes clinical work through workflow, professional roles, communication routines and organisational arrangements, not only through model performance (4). When AI produces, summarises, simplifies or restructures radiology text, it enters a communicative layer of care in which phrasing, certainty, emphasis and omission may influence clinical judgement.

In this study, the term AI-generated radiology narrative refers to written text created or assisted by AI that describes imaging findings, drafts or organises an impression, summarises an existing report, simplifies a report for patients or highlights information relevant to clinical decision-making. These outputs differ from conventional computer-aided detection because they affect the language through which imaging evidence is interpreted and transferred between professionals. A detection algorithm may flag a pulmonary nodule, whereas a narrative system may describe its likely significance, place it within a differential diagnosis, express uncertainty and suggest future action. Since report quality and structured communication directly affect patient care, AI-generated narratives introduce a new layer of authorship, interpretation and accountability into clinical reporting (5).

The possible benefits of AI-generated radiology narratives are substantial. They may reduce documentation burden, support clinicians who need to interpret lengthy reports quickly, standardise repetitive sections, increase visibility of follow-up recommendations and translate technical language into patient-friendly explanations. However, large language model applications in radiology have raised concerns regarding hallucination, factual incompleteness, omitted caveats, overconfident phrasing and weak evaluation methods (6). Alabed et al. reported that large-language-model simplification improved readability and perceived understanding of radiology reports, but the same evidence base identified errors that could be clinically significant (7). In clinical reporting, Huang et al. found efficiency gains with generative AI-assisted radiology reporting without a corresponding measurable improvement in report quality, indicating that speed alone cannot be treated as a sufficient safety marker (8).

Clinician trust is therefore central to the safe implementation of AI-generated radiology narratives. If clinicians do not trust these outputs, potentially useful tools may fail to reduce workload or improve communication. If clinicians over-trust fluent AI-generated language, unsafe automation bias may occur. The appropriate implementation goal is not maximum trust, but calibrated reliance, in which clinicians use AI-generated narratives when the evidence, context, workflow and safeguards support safe use, and question them when clinical risk is high, traceability is weak or accountability is unclear. Evidence that AI assistance affects radiologists differently across clinicians and tasks further supports the need to understand trust as situated, conditional and shaped by clinical context rather than as a fixed attitude toward technology (9).

Existing research on AI in radiology has largely prioritised technical performance. Image-recognition studies show that AI can perform strongly in defined diagnostic tasks; for example, McKinney et al. evaluated an AI system for breast cancer screening and demonstrated high diagnostic performance under controlled conditions (10). Lång et al. similarly showed the promise of AI-supported screen reading in mammography screening (11). These studies are important because they demonstrate the potential of AI to support detection, prioritisation and consistency, yet radiology practice cannot be reduced to classification. Reporting requires comparison with previous imaging, interpretation of clinical history, management of uncertainty, communication with referrers and professional responsibility for the wording of the final impression.

A second strand of literature concerns regulation, commercial availability and scientific evidence for radiology AI. Regulatory authorisation indicates that a device has met requirements for a specified intended use, but it does not guarantee clinical effectiveness across every hospital, scanner environment, patient group or reporting convention. Muehlematter, Daniore and Vokinger argued that approval pathways for AI and machine-learning medical devices require careful interpretation in relation to real-world clinical use (12). Van Leeuwen et al. found variability in the strength of published scientific evidence supporting commercially available radiology AI products (13). For AI-generated narratives, this distinction is especially important because the output is not merely a probability score but a professionally meaningful message that may be acted on directly. Local validation is therefore not only a regulatory or administrative requirement but a condition of trustworthy clinical communication.

Large language models have expanded the narrative functions of radiology AI. Earlier automated report-generation systems generally attempted to generate reports from images, whereas newer systems may extract information from reports, convert free text into structured language, simplify technical wording, produce summaries or draft impressions. Reichenpfader, Muller and Denecke described the expanding use of large language models for information extraction from radiology reports (14). These functions carry different levels of clinical risk. A patient-friendly explanation based on a radiologist-signed report may be lower risk than an AI-generated diagnostic impression based directly on imaging, but even simplification can become unsafe if it removes uncertainty, comparison information, limitations or follow-up conditions. Sunshine et al. therefore emphasised that understandability must be assessed alongside quality when AI produces radiology summaries (15).

Evaluation of radiology narratives is difficult because surface fluency is not equivalent to clinical fidelity. Traditional natural-language metrics may reward smooth wording or similarity to reference reports while failing to detect clinically important omissions. Agarwal et al. proposed clinical evaluation and quality scoring for radiology report generation to address this limitation (16). Gaur et al. also argued that evaluation must distinguish linguistic quality from clinical correctness (17). This distinction is central to clinician trust because clinicians do not rely on reports because they are grammatical; they rely on them because they preserve decisive findings, uncertainty, limitations and recommended actions. A polished AI-generated narrative may still be unsafe if it omits a secondary abnormality, weakens a caveat or turns a cautious impression into a definitive statement.

Trust in healthcare AI is shaped by perceived accuracy, transparency, usability, education, accountability and institutional endorsement. Asan, Bayrak and Choudhury reviewed human trust in healthcare AI and showed that trust is a multidimensional relationship between users, systems and settings (18). Broader AI trust research also indicates that users may either overvalue algorithmic recommendations or reject algorithms after visible errors, depending on perceived authority, prior experience and task context (19). In clinical care, these tendencies have heightened consequences because clinicians remain professionally responsible for patient outcomes even when AI participates in the information environment.

For this reason, calibrated reliance is more appropriate than general trust-building. Human-AI interaction research suggests that cognitive forcing functions may reduce overreliance by prompting users to pause before accepting automated advice (20). Interpretability research also shows that explanations may create false confidence if users treat them as proof of correctness rather than as limited decision aids (21). Clinicians therefore need task-relevant transparency rather than abstract technical explanations. Tonekaboni et al. found that clinicians value explanations that are contextual and useful for clinical work (22). In radiology narratives, this means that useful transparency includes the source of the text, the evidence supporting the impression, the uncertainty that remains, whether the narrative was derived from a signed report or direct image interpretation, and whether a radiologist has verified the output.

Bias and equity further complicate trust. Imaging AI systems may perform differently across populations, institutions and acquisition contexts. Gichoya et al. showed that AI systems can recognise patient race from medical images despite mechanisms that are not easily explained by human observers (23). Seyyed-Kalantari et al. demonstrated underdiagnosis bias when chest radiograph algorithms were applied to underserved patient populations (24). These findings challenge the assumption that imaging AI is inherently objective. Narrative systems may reproduce or amplify bias through reporting language, omitted context, weakened performance in under-represented settings or differential handling of uncommon but high-consequence findings. Trust must therefore include local validation across patient groups, scanners, protocols and clinical pathways.

Governance literature frames healthcare AI as an ethical and socio-technical intervention. The World Health Organization emphasises autonomy, safety, transparency, responsibility, inclusiveness and sustainability in health AI (25). Radiology ethics guidance similarly stresses human control, fairness, reliability and accountability (26). The European Society of Radiology has also considered how risk management and traceability requirements under the AI Act apply to radiological practice (27). For reporting, these principles translate into practical obligations: AI-generated text should be labelled, source evidence should remain accessible, oversight should be documented, local validation should be recorded, audit trails should exist and error feedback routes should be available.

Implementation research explains why technically impressive AI may still fail in practice. Kelly et al. identified evaluation, regulation, human factors and deployment as key challenges for achieving clinical impact with AI (28). Beede et al. demonstrated that even high-performing deep-learning systems can be reshaped by workflow, staffing, trust and patient interaction after deployment (29). These lessons apply directly to AI-generated radiology narratives. If AI-generated text creates additional screens, unclear authorship, poor integration or uncertainty about who owns the final report, clinicians may ignore it, copy it uncritically or use it inconsistently. Trust is therefore produced through the relationship between the tool, the task, the clinician and the organisation.

Despite growing evidence on radiology AI performance, regulation and implementation, less is known about how clinicians themselves judge AI-generated radiology narratives as clinical communication tools. This gap is important because trust in narrative AI is not only a technical judgement about model accuracy; it is also an interpretive judgement about uncertainty, provenance, professional responsibility, workflow fit and the consequences of acting on generated language. A qualitative approach is appropriate for examining these situated judgements because it allows clinicians to describe how trust is negotiated in relation to specific tasks, risks and safeguards. This study therefore aimed to explore how clinicians judge the usefulness, limits and required safeguards of AI-generated radiology narratives when these narratives may inform patient-care decisions.

MATERIALS AND METHODS

This study used a qualitative descriptive design informed by interpretive inquiry to explore clinician trust in AI-generated radiology narratives for patient-care decisions. A qualitative approach was appropriate because the research question concerned professional judgement, perceived risk, accountability, workflow experience, uncertainty and meaning-making rather than measurement of a fixed attitude or estimation of prevalence. The study was designed to examine how clinicians interpreted AI-generated narrative outputs across different clinical contexts and how they described the conditions under which such outputs might be considered useful, unsafe, verifiable or professionally acceptable. The methodological orientation was consistent with qualitative inquiry, in which participants' accounts are analysed within their clinical and organisational context to generate an explanatory understanding of the phenomenon under study (30).

The study population consisted of clinicians whose work involved producing, interpreting or acting on radiology reports in patient care. Purposive sampling was used to obtain variation in clinical role, proximity to radiological interpretation and downstream dependence on imaging communication. The participant group comprised ten clinicians, including a consultant radiologist, radiology registrar, emergency physician, oncologist, respiratory physician, orthopaedic surgeon, neurologist, intensive care consultant, cardiologist and general practitioner with imaging referral responsibilities. This sampling approach was intended to include both radiology professionals who understand report production and non-radiology clinicians who use radiology narratives to guide clinical decisions. The sample was not intended to be statistically representative; rather, it was selected to support depth, diversity of clinical perspective and comparison across decision contexts in which AI-generated radiology narratives may carry different levels of risk.

Participants were recruited because of their professional experience with radiology reports and their relevance to the clinical use of imaging narratives. Eligibility required current or recent clinical practice in a specialty where radiology reports contribute to diagnostic, therapeutic, referral, discharge, follow-up or patient communication decisions. Clinicians were eligible if they had experience using radiology reports in routine patient care and were able to discuss potential uses, risks and safeguards of AI-generated radiology text. Clinicians were not included if their role did not involve direct use, interpretation or communication of radiology report information for patient-care decisions. Written informed consent was obtained from all participants before participation, and the study was conducted after ethical approval from Clinical Medicine, Prima Indonesia University, Indonesia.

Data were collected using semi-structured interviews. The interview format allowed the discussion to remain focused on the study aim while giving participants sufficient flexibility to describe their own clinical reasoning, workflow experiences and concerns about AI-generated narratives. The interview guide moved from general experience with radiology reports to more specific questions about AI-generated summaries, draft impressions, patient-friendly explanations, uncertainty language, traceability, radiologist supervision, workflow integration, accountability and safeguards required for safe adoption. This sequence was used to ground responses in actual clinical reporting practices before asking participants to reflect on AI-generated text. Open-ended questions were used to avoid forcing participants into a simple positive or negative position toward AI and to elicit conditional judgements about when AI-generated narratives might be useful, risky or unacceptable.

The phenomenon of interest was clinician trust in AI-generated radiology narratives. For the purposes of the study, AI-generated radiology narratives were defined as AI-assisted or AI-created written outputs that describe imaging findings, organise or draft impressions, summarise existing reports, simplify reports for patients or highlight report information relevant to decision-making. The primary analytic focus was not whether participants supported or opposed AI in general, but how they judged trust in relation to clinical risk, narrative function, source visibility, uncertainty, radiologist oversight, workflow fit and accountability. These domains were treated as sensitising concepts during interview development while allowing themes to be developed from participant accounts.

Interview transcripts were analysed using reflexive thematic analysis. The analytic process followed familiarisation with the transcripts, generation of initial codes, development of candidate themes, review of themes against the dataset, refinement of theme definitions and analytic writing. This approach was selected because reflexive thematic analysis is suitable for identifying patterned meanings across qualitative data while recognising the active interpretive role of the researcher in theme development (31). Coding was conducted at both semantic and interpretive levels. Semantic coding captured explicit statements about labelling, source linkage, uncertainty language, validation, radiologist verification, workflow integration and error reporting. Interpretive coding examined the professional assumptions underlying these statements, including accountability anxiety, concern about fluent error, interprofessional dependence, conditional reliance and the need for organisational governance.

Theme development was iterative and comparative. Codes were examined within and across participant roles to identify areas of convergence, contrast and clinical nuance. Particular attention was given to differences between radiology participants and downstream report users, as well as to distinctions between low-risk narrative functions and high-consequence decisions such as cancer staging, emergency discharge, surgical referral and critical-care escalation. Contrasting cases were actively considered so that the analysis did not reduce clinicians to being simply trusting or distrusting of AI. For example, accounts in which participants valued AI summaries for reducing cognitive load were examined alongside statements limiting reliance when uncertainty, provenance or accountability was unclear. This process supported the development of themes that represented trust as conditional, task-specific and shaped by organisational safeguards rather than as a fixed attitude toward technology.

Trustworthiness was addressed through transparent linkage between the research question, interview domains, coding process, theme development and final interpretation. Credibility was supported by using participants with direct clinical experience of radiology report use, by grounding analysis in role-specific accounts and by retaining contrasting perspectives during theme refinement. Dependability was supported through a documented analytic process that moved from transcript familiarisation to coding, theme review and final thematic definition. Confirmability was strengthened by maintaining an audit trail of analytic decisions and by distinguishing participant claims from the interpretive meanings developed through thematic analysis. Transferability was supported by purposive inclusion of clinicians from radiology and multiple non-radiology specialties, enabling readers to judge the relevance of the findings to other settings where AI-generated radiology narratives may influence patient-care decisions.

The analysis generated five final themes: narrative usefulness as cognitive scaffolding rather than replacement; traceability and validation as the foundation of trust; accountability anxiety in higher-risk decisions; workflow fit and interprofessional communication; and provisional trust shaped by governance and learning. These themes were developed to capture how clinicians described the conditions under which AI-generated radiology narratives could support clinical communication while remaining bounded by uncertainty, human oversight and organisational responsibility. The findings were reported using representative participant quotations to illustrate the analytic meaning of each theme and to show how trust was negotiated across clinical roles and decision contexts.

Ethical conduct was maintained throughout the study. Participation was voluntary, informed consent was obtained, and confidentiality was protected in the reporting of participant accounts. Participant identifiers were used in place of names, and quotations were presented in a way that preserved clinical meaning while avoiding direct personal identification. The study complied with institutional ethical requirements, and the data were handled in accordance with the consent process and confidentiality obligations for qualitative interview material.

RESULTS

The analysis included semi-structured interview transcripts from ten clinicians whose work involved producing, interpreting or acting on radiology reports in patient care. Participants represented both radiology-facing and downstream clinical roles, allowing comparison between those responsible for report generation or verification and those using radiology narratives to guide treatment, referral, discharge, escalation and patient communication decisions.

Table 1. Participant Characteristics

Participant ID	Clinical Role or Specialty	Relationship to Radiology Narratives	Main Decision Context
P1	Consultant radiologist	Report production and verification	Diagnostic interpretation, report sign-off, source verification
P2	Emergency physician	Downstream report user	Acute decision-making, discharge, escalation
P3	Oncologist	Downstream report user	Staging, treatment planning, follow-up decisions
P4	Respiratory physician	Downstream report user	Diagnostic correlation, follow-up planning
P5	Orthopaedic surgeon	Downstream report user	Surgical referral, imaging-based management
P6	Neurologist	Downstream report user	Diagnostic clarification, treatment direction
P7	Intensive care consultant	Downstream report user	Critical-care escalation, urgent interpretation
P8	Radiology registrar	Report production and review	Draft review, wording accuracy, uncertainty preservation
P9	Cardiologist	Downstream report user	Risk assessment, diagnostic confirmation
P10	General practitioner	Referrer and downstream report user	Referral decisions, patient explanation, follow-up coordination

The sample included clinicians from radiology, emergency medicine, oncology, respiratory medicine, orthopaedics, neurology, intensive care, cardiology and general practice. This variation supported analysis of trust as a conditional judgement shaped by clinical role, decision consequence, workflow dependence and perceived accountability.

The participant group reflected the clinical pathway through which radiology narratives move from specialist interpretation to action by referring and treating clinicians. Radiology participants focused strongly on source verification, authorship and preservation of uncertainty, whereas downstream clinicians emphasised actionable summaries, rapid interpretation, workflow integration and access to clarification. Across roles, participants did not describe trust in AI-generated radiology narratives as simple acceptance or rejection. Instead, trust was described as provisional, task-specific and dependent on whether the AI-generated text remained traceable, clinically cautious, human-supervised and compatible with existing communication systems.

Thematic analysis generated five themes: narrative usefulness as cognitive scaffolding rather than replacement; traceability and validation as the foundation of trust; accountability anxiety in higher-risk decisions; workflow fit and interprofessional communication; and provisional trust shaped by governance and learning. These themes described how clinicians judged AI-generated radiology narratives in relation to the clinical function of the text, the level of risk attached to the decision, the visibility of the source evidence, the availability of radiologist oversight and the presence of organisational safeguards.

Table 2. Themes, Analytic Meaning, and Representative Quotations

Theme	Analytic Meaning	Representative Quotation
Narrative usefulness as cognitive scaffolding rather than replacement	AI-generated narratives were valued when they organised complex reports, highlighted key findings and reduced cognitive load without replacing clinician judgement.	"Radiology reports can be very long; a structured AI summary would help, but it must not make uncertainty disappear."
Narrative usefulness as cognitive scaffolding rather than replacement	Clinicians valued speed only when the narrative preserved clinical nuance and avoided unsafe closure.	"I need the bottom line quickly, but a confident wrong bottom line is dangerous."
Traceability and validation as the foundation of trust	Trust depended on knowing where the narrative came from and whether it could be checked against source evidence.	"I would not sign it unless I could verify where the narrative came from."
Traceability and validation as the foundation of trust	Labelling and uncertainty language were considered essential safeguards for AI-generated sections.	"The wording can look very polished, and that polish can hide what has been missed."
Accountability anxiety in higher-risk decisions	Participants were cautious when AI-generated wording could influence treatment, discharge, staging, surgery or escalation decisions.	"I might use it for routine follow-up summaries, but not for cancer staging without direct radiologist confirmation."
Workflow fit and interprofessional communication	Trust was strengthened when AI narratives were embedded in existing clinical systems and preserved routes for clarification.	"If it is in the PACS and shows the changes from the previous scan, I would use it; if it needs another login, people will ignore it."
Workflow fit and interprofessional communication	Clinicians distinguished technical output from professional communication and valued the ability to question or clarify findings.	"I trust my radiology colleagues because I can call them. I cannot call an algorithm unless there is a feedback route."
Provisional trust shaped by governance and learning	Trust was treated as organisational rather than purely individual, requiring policy, audit, feedback and accountability structures.	"The issue is not just the AI sentence, but whether the organisation has checked it, labelled it and created a route for correction."

The first theme, narrative usefulness as cognitive scaffolding rather than replacement, described clinicians' willingness to use AI-generated narratives as aids to attention and interpretation. Participants considered AI potentially useful for organising lengthy radiology reports, drawing attention to key findings and helping busy clinicians identify clinically relevant information more efficiently. This usefulness was especially apparent for complex reports with long descriptive sections, follow-up comparisons or multiple findings. P5 expressed this conditional value clearly, stating, "Radiology reports can be very long; a structured AI summary would help, but it must not make uncertainty disappear." This quotation captured the central boundary of the theme: AI-generated narrative support was acceptable when it improved structure and readability, but unsafe when it simplified uncertainty into unjustified certainty.

Emergency and acute-care perspectives reinforced the distinction between useful speed and unsafe closure. P2 stated, "I need the bottom line quickly, but a confident wrong bottom line is dangerous." This account showed that participants did not reject concise AI-generated summaries; rather, they rejected summaries that could create premature certainty or suppress clinical caveats. The value of AI-generated text was therefore not its ability to produce fluent language, but its ability to support clinical attention while preserving the interpretive caution already embedded in radiology reporting. Participants treated AI-generated narratives as potentially helpful cognitive scaffolds, not as independent clinical authorities.

Radiology participants were particularly sensitive to the distinction between fluency and correctness. P8 warned, “The wording can look very polished, and that polish can hide what has been missed.” This concern suggested that polished AI-generated language may increase perceived authority even when the content is incomplete, unsupported or insufficiently cautious. Participants therefore described narrative fluency as double-edged. It could make reports easier to use, but it could also obscure omissions or exaggerate certainty. The theme showed that clinicians were willing to use AI-generated narratives when they supported attention, organisation and communication, but not when they replaced verification, contextual judgement or professional responsibility.

The second theme, traceability and validation as the foundation of trust, described participants’ emphasis on source visibility and local verification. Clinicians wanted to know whether an AI-generated narrative was derived from image data, a radiologist-verified report, structured fields, prior imaging, a reporting template or model-generated inference. P1 stated, “I would not sign it unless I could verify where the narrative came from.” This response showed that trust was not built through general claims about AI performance alone. Instead, clinicians wanted to trace the specific narrative in the specific case back to its evidential source.

Participants also emphasised that AI-generated sections should be clearly labelled and should preserve uncertainty where uncertainty remained. Labelling was not viewed as a superficial formatting issue but as a safety requirement that allowed clinicians to distinguish radiologist-authored conclusions from AI-assisted wording. Source linkage, visible uncertainty language and local validation were repeatedly treated as conditions for responsible use. For radiology participants, traceability was closely linked to sign-off and professional accountability. For downstream clinicians, traceability was linked to whether they could safely act on a summary, impression or patient-friendly explanation. Across roles, trust depended on whether AI-generated text could be checked, challenged and validated within the clinical pathway.

The third theme, accountability anxiety in higher-risk decisions, captured participants’ caution when AI-generated narratives could influence consequential clinical decisions. Clinicians were more willing to use AI for routine summaries, administrative preparation, patient-friendly explanations based on signed reports and low-risk follow-up planning. They were less willing to rely on AI-generated wording when small changes in phrasing could alter treatment, discharge, staging, surgical referral or escalation decisions. P3 stated, “I might use it for routine follow-up summaries, but not for cancer staging without direct radiologist confirmation.” This distinction demonstrated that participants judged trust according to the consequence of error rather than according to the presence of AI alone.

Accountability anxiety was not merely concern about blame. Participants described a deeper professional responsibility for decisions made using radiology information. In high-risk settings, terms such as stable, suspicious, likely, cannot exclude, progressive or correlate clinically may determine whether a patient is discharged, investigated further, referred for surgery, escalated to critical care or started on oncological treatment. Participants therefore worried that AI-generated narratives could shift clinical interpretation if they changed the force, certainty or emphasis of radiology language. This theme showed that trust was strongest when AI supported lower-risk communication and weakest when it appeared to generate or reshape conclusions that could directly change management without explicit human oversight.

The fourth theme, workflow fit and interprofessional communication, described trust as dependent on integration into existing clinical systems and communication pathways. Participants were more likely to trust AI-generated narratives if they appeared within familiar reporting environments, preserved access to prior imaging and allowed clinicians to identify whether a radiologist had reviewed the output. P7 stated, “If it is in the PACS and shows the changes from the previous scan, I would use it; if it needs another login, people will ignore it.” This account showed that technical usefulness could be undermined by poor workflow design. Even a potentially accurate AI narrative might be disregarded if it required separate access, duplicated work or appeared outside the trusted reporting pathway.

Participants also distinguished AI-generated output from relational professional trust. P10 stated, "I trust my radiology colleagues because I can call them. I cannot call an algorithm unless there is a feedback route." This quotation showed that trust in radiology communication is not limited to written text; it is supported by the ability to ask questions, clarify uncertainty and resolve disagreement. For AI-generated narratives to become trustworthy within clinical work, participants expected visible responsibility and a route for feedback or correction. Workflow fit therefore included not only software integration but also interprofessional communication, escalation pathways and the ability to respond when AI-generated text was unclear or potentially wrong.

The fifth theme, provisional trust shaped by governance and learning, described trust as an organisational process that must be monitored over time. Participants wanted hospital-level policies defining which AI-generated narrative functions were permitted, who could approve them, how errors would be reported, how outputs would be audited and how model updates would be monitored. P4 stated, "The issue is not just the AI sentence, but whether the organisation has checked it, labelled it and created a route for correction." This theme showed that participants did not expect individual clinicians to carry the full burden of judging AI safety in isolation. They wanted institutional structures that made safe use possible.

Governance was especially important because participants viewed AI-generated radiology narratives as dynamic outputs that could change with software updates, local reporting practices, scanner protocols and clinical pathways. Trust was therefore described as renewable rather than permanent. Initial validation was not considered sufficient if the model later changed or if the tool was applied to new clinical settings. Participants expected audit trails, feedback channels, radiologist oversight, user training and defined accountability pathways. In this theme, trust was provisional because it depended on ongoing evidence that the AI-generated narrative remained accurate, useful, labelled, traceable and clinically safe.

Table 3. Conditional Trust Framework Derived from Participant Accounts

Trust Condition	Low-Risk Narrative Use	Higher-Risk Narrative Use	Required Safeguard
Narrative function	Summary of signed report	Draft diagnostic impression	Human verification
Clinical consequence	Routine follow-up support	Cancer staging or emergency discharge	Radiologist confirmation
Source visibility	Linked to verified report	Directly generated or unclear source	Source traceability
Uncertainty handling	Preserves caveats	Converts uncertainty into certainty	Explicit uncertainty language
Workflow integration	Embedded in PACS or reporting system	Separate interface or copied text	Integration with clinical systems
Accountability	Clear reviewer or sign-off pathway	Unclear authorship	Defined responsibility
Error correction	Feedback route available	No correction mechanism	Audit and reporting pathway
Institutional oversight	Locally approved use	Unvalidated deployment	Governance and monitoring

The conditional trust framework demonstrated that participants did not assess AI-generated radiology narratives through a single criterion. They weighed narrative function, clinical consequence, source visibility, uncertainty, workflow integration, accountability and governance simultaneously. AI-generated summaries based on signed reports were viewed as more acceptable when they preserved caveats and remained linked to the source report. AI-generated impressions that could alter treatment were viewed as higher risk and required explicit radiologist confirmation. Trust was therefore strongest when AI-generated narratives were low-risk, traceable, integrated and reviewable, and weakest when they were high-consequence, unlabelled, detached from source evidence or unsupported by clear accountability.

Overall, the findings showed that clinicians viewed AI-generated radiology narratives as socio-technical communication tools rather than independent clinical authorities. Participants recognised potential benefits in reducing cognitive load, improving report navigation and supporting patient-facing communication, but they consistently placed boundaries around acceptable use. Trust was calibrated through clinical risk, traceability, uncertainty preservation, radiologist oversight, workflow fit and organisational governance. The analysis therefore indicated that safe implementation should not aim to make clinicians broadly trust AI-generated radiology narratives, but to support limited, transparent and accountable reliance in contexts where the narrative function, source evidence and decision consequence are clearly defined.

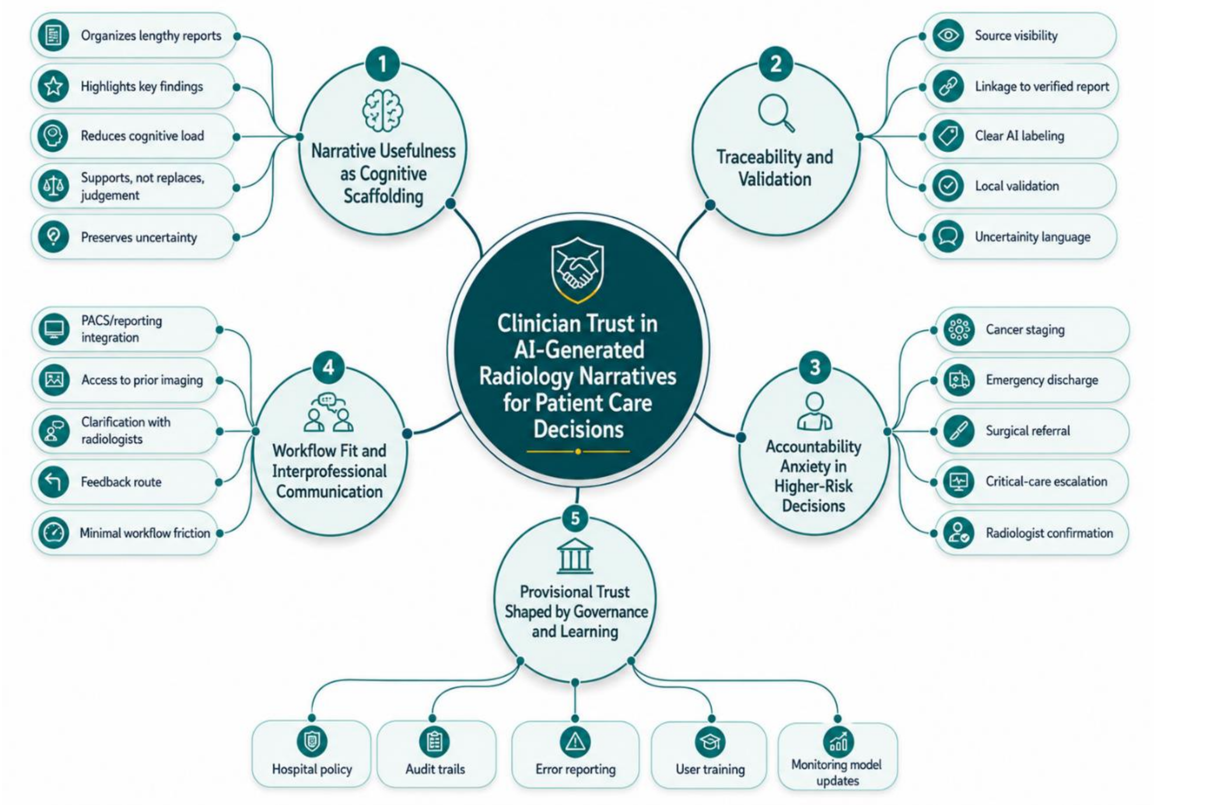

Figure 1 Thematic network map of clinician trust in AI-generated radiology narratives for patient-care decisions. The conceptual framework illustrates clinician trust as a conditional and socio-technical construct shaped by five interrelated qualitative themes: narrative usefulness as cognitive scaffolding, traceability and validation, accountability anxiety in higher-risk decisions, workflow fit and interprofessional communication, and provisional trust shaped by governance and learning. Subthemes show that trust depended on whether AI-generated narratives organized information without replacing judgement, preserved uncertainty, remained linked to verified source reports, supported radiologist confirmation in high-risk decisions, integrated into existing PACS/reporting workflows, and operated within institutional safeguards such as audit trails, error reporting, user training, and monitoring of model updates.

DISCUSSION

This qualitative study explored how clinicians judged trust in AI-generated radiology narratives when such narratives could inform patient-care decisions. The findings showed that clinicians did not understand trust as simple acceptance or rejection of AI-generated text. Instead, trust was constructed as a conditional, task-specific and professionally accountable judgement shaped by the function of the narrative, the clinical consequence of error, the visibility of source evidence, the preservation of uncertainty, the availability of radiologist oversight, the fit with existing workflow and the presence of organisational governance. Across the five themes, AI-generated radiology narratives were viewed as potentially useful communication aids, but not as independent clinical authorities.

The theme of narrative usefulness as cognitive scaffolding provides an important refinement to the common efficiency argument for AI adoption in radiology. Participants recognised that AI-generated summaries or structured narratives could help clinicians navigate long reports, identify key findings and reduce cognitive load during busy clinical workflows. This potential value is consistent with the broader pressures facing radiology services, where increasing imaging demand and workforce limitations create a need for tools that improve communication efficiency without weakening diagnostic safety (1). However, participants made a clear distinction between helpful structuring and unsafe simplification. They valued AI-generated narratives when these outputs organised information and supported attention, but they were concerned when concise or fluent wording risked removing caveats, comparison details or uncertainty. This finding is particularly important because radiology reports often carry clinical meaning through careful wording. Terms such as possible, likely, cannot exclude, stable, suspicious or correlate clinically may influence downstream management. If AI-generated text improves readability while weakening these interpretive signals, the apparent communication gain may create clinical risk.

The findings also reinforce the distinction between linguistic fluency and clinical fidelity. Participants were concerned that polished AI-generated language could conceal omissions, unsupported conclusions or excessive certainty. This concern aligns with current evidence on large language model use in radiology, where readability and user satisfaction may improve while clinically meaningful errors remain possible (6,7). In radiology reporting, a narrative cannot be judged only by grammatical quality or surface coherence. Its safety depends on whether it preserves decisive findings, uncertainty, limitations, comparison with prior imaging and recommended actions. The present findings suggest that clinicians are aware of the persuasive force of fluent AI text and that trust depends on mechanisms that make the narrative verifiable rather than merely readable. This supports the need for evaluation approaches that assess management impact, omission detection, uncertainty preservation and clinical correctness rather than relying on language quality alone (16,17).

Traceability and validation emerged as central foundations of trust. Clinicians wanted to know whether an AI-generated narrative came from image data, a signed radiology report, structured reporting fields, prior imaging, a template or unsupported generative inference. This form of traceability functioned as a practical version of explainability. Participants did not request abstract model architecture explanations; they wanted clinically useful transparency that allowed them to verify the source and status of the narrative in the specific case. This finding is consistent with evidence that clinicians prefer explanations that are contextual and useful for clinical work rather than technically exhaustive (22). For AI-generated radiology narratives, useful transparency means visible labelling, source linkage, uncertainty language, auditability and clear indication of whether a radiologist has reviewed or signed the output. Trust was therefore not based on the general reputation of AI but on the ability to trace and validate the particular text being used for a particular patient-care decision.

Accountability anxiety was especially prominent in higher-risk decisions. Participants were more willing to use AI-generated text for routine summaries, follow-up organisation or patient-friendly explanations derived from verified reports, but they were cautious when narrative outputs could influence cancer staging, emergency discharge, surgical referral, critical-care escalation or other high-consequence decisions. This finding shows that clinician trust is not only a matter of perceived accuracy; it is also shaped by professional responsibility. Clinicians remain accountable for decisions made in patient care even when AI contributes to the information environment. As a result, their willingness to rely on AI-generated narratives was closely tied to whether responsibility, verification and escalation pathways were clear. This supports the argument that AI-generated impressions or management-relevant narrative outputs should remain under explicit human oversight, particularly when wording may directly alter treatment, referral or discharge decisions.

Workflow fit and interprofessional communication further showed that trust is socio-technical rather than purely technical. Participants were more likely to trust AI-generated narratives when they were embedded within existing PACS, reporting systems or electronic health record pathways and when they preserved access to prior imaging and routes for clarification. Poor integration, separate logins or unclear document status were seen as barriers to use even when the concept of AI support was acceptable. This aligns with implementation research showing that AI tools may fail to produce clinical impact when deployment does not account for workflow, staffing, communication and user interaction (28,29). In radiology, trust is also relational. Clinicians trust reports not only because of their content but because they are produced within a professional system in which radiologists can be contacted, questioned and held accountable. AI-generated narratives therefore need feedback routes, escalation channels and clear ownership if they are to function safely within clinical communication.

The theme of provisional trust shaped by governance and learning indicates that participants viewed safe AI use as an organisational responsibility. They expected hospital-level policies, local validation, defined sign-off procedures, audit trails, error reporting, user training and monitoring of model updates. This finding is consistent with ethical and governance frameworks that emphasise transparency, responsibility, safety, fairness and human oversight in healthcare AI (25,26). It also reflects the lifecycle nature of AI safety. A narrative system that performs acceptably during initial evaluation may change after software updates, workflow modifications, new reporting conventions or application to different patient groups. Trust therefore needs to be renewable and continuously monitored rather than granted permanently at the point of adoption. This is particularly relevant in radiology, where variation in scanners, protocols, reporting styles and patient populations may affect both technical performance and narrative interpretation.

The findings also have implications for bias and equity. Although participants focused more directly on traceability, accountability and workflow, their concerns about local validation are closely connected to fairness. Imaging AI can perform differently across patient populations and institutional contexts, and prior research has shown that AI systems may detect sensitive patient characteristics from medical images or underdiagnose conditions in underserved populations (23,24). AI-generated narratives may reproduce such problems through differential accuracy, selective omission, overconfident phrasing or failure to preserve clinically relevant context. Local validation should therefore examine not only overall narrative quality but also performance across patient groups, imaging protocols, specialties and high-risk clinical scenarios. Trust in AI-generated radiology narratives should include evidence that the system performs safely and consistently for the population in which it is deployed.

This study contributes to the literature by shifting attention from whether clinicians trust AI in general to how they calibrate trust in a specific communicative use of AI. AI-generated radiology narratives occupy a distinctive position because they translate imaging evidence into language that may influence clinical action. They are neither purely diagnostic algorithms nor ordinary administrative text. Their safety depends on how accurately they preserve clinical meaning, how visibly they represent uncertainty, how clearly they identify their source and how responsibly they are integrated into professional workflows. The thematic network developed in this study suggests that trust is strongest when narratives are used as cognitive scaffolds, linked to verified sources, embedded in established systems, supported by radiologist oversight and governed through institutional monitoring. Trust is weakest when narratives appear authoritative, unlabelled, detached from source evidence or capable of altering high-risk decisions without human confirmation.

Several limitations should be considered when interpreting these findings. The study included ten clinicians selected purposively to provide variation across radiology and downstream clinical roles. This sample allowed in-depth exploration of clinically diverse perspectives, but the findings are not intended to be statistically generalisable. Transferability depends on the similarity of other clinical settings, radiology workflows, AI governance structures and clinician experience with AI-generated text. The study also relied on clinician accounts of trust, perceived usefulness and anticipated safeguards rather than direct observation of long-term AI deployment in routine radiology practice. Participants' views may therefore reflect both actual experience with radiology communication and prospective judgement about emerging AI tools. In addition, the findings may be shaped by the specialties represented in the sample; other groups such as paediatrics, pathology-linked oncology teams, radiographers, patients, hospital managers and medico-legal stakeholders may identify additional concerns. Future research should examine AI-generated radiology narratives in real-world clinical workflows, compare specialty-specific patterns of reliance, evaluate patient-safety outcomes, and explore how patients interpret AI-simplified radiology explanations.

The practical implication of this study is that implementation should begin with low-risk, auditable and clearly labelled narrative functions rather than high-consequence autonomous impressions. AI-generated summaries based on radiologist-signed reports, patient-friendly explanations linked to the original report and administrative organisation of reporting text may be more acceptable starting points than unverified diagnostic conclusions. Higher-risk uses should require stronger verification, explicit radiologist confirmation and documented accountability. Clinician training should not only teach users how to operate AI interfaces but also how to challenge fluent AI language, identify missing uncertainty, verify source evidence and escalate concerns. In this sense, the goal of implementation should not be to increase clinician trust broadly, but to support calibrated reliance: appropriate use when safeguards are present and active scepticism when risk, traceability or accountability is unclear.

CONCLUSION

This qualitative study found that clinician trust in AI-generated radiology narratives is conditional, provisional and shaped by the clinical context in which narrative text is used. Clinicians viewed AI-generated narratives as potentially useful cognitive scaffolds that may organise lengthy reports, highlight key findings, reduce cognitive load and support communication, but they did not regard them as substitutes for radiologist judgement or professional accountability. Trust depended on clear labelling, linkage to verified source reports, preservation of uncertainty, local validation, workflow integration, access to clarification, radiologist confirmation for high-risk decisions and organisational safeguards such as audit trails, error reporting, user training and monitoring of model updates. AI-generated radiology narratives should therefore be implemented as transparent, traceable and accountable communication tools rather than independent clinical authorities. Safe adoption should begin with lower-risk, reviewable functions and should require stronger human oversight when narrative outputs may influence cancer staging, emergency discharge, surgical referral, critical-care escalation or other high-consequence decisions. The findings support a model of calibrated reliance in which clinicians use AI-generated narratives when they are source-linked, contextually appropriate and governed, while retaining the ability and responsibility to question, verify and escalate uncertainty.

REFERENCES

1. Royal College of Radiologists. Clinical radiology workforce census 2024. London: Royal College of Radiologists; 2025. Available from: https://www.rcr.ac.uk/news-policy/policy-reports-initiatives/clinical-radiology-census-reports/

2. Singh R, et al. How AI is used in FDA-authorized medical devices. NPJ Digit Med. 2025. Available from: https://www.nature.com/articles/s41746-025-01800-1

3. European Society of Radiology. What the radiologist should know about artificial intelligence. Insights Imaging. 2023. Available from: https://doi.org/10.1186/s13244-023-01435-4

4. Chou SHS, et al. Integrating artificial intelligence into radiology workflow: deployment lessons. Radiology. 2022. Available from: https://pubs.rsna.org/

5. Bae MS, et al. Radiology report quality, structured reporting and communication in clinical care. Korean J Radiol. 2021. Available from: https://doi.org/10.3348/kjr.2020.0946

6. Keshavarz P, et al. ChatGPT in radiology: a systematic review of performance, pitfalls and future directions. Clin Imaging. 2024. Available from: https://www.sciencedirect.com/science/article/pii/S2211568424001050

7. Alabed S, et al. Large language models for simplifying radiology reports: systematic review and meta-analysis. Lancet Digit Health. 2026. Available from: https://www.thelancet.com/journals/landig/article/PIIS2589-7500(25)00142-6/fulltext

8. Huang J, et al. Efficiency and quality of generative AI-assisted radiology reporting in clinical care. JAMA Netw Open. 2025. Available from: https://jamanetwork.com/journals/jamanetworkopen/fullarticle/2834943

9. Yu F, et al. Heterogeneity and predictors of the effects of AI assistance on radiologists. Nat Med. 2024. Available from: https://www.nature.com/articles/s41591-024-02850-w

10. McKinney SM, Sieniek M, Godbole V, Godwin J, Antropova N, Ashrafian H, et al. International evaluation of an AI system for breast cancer screening. Nature. 2020;577:89-94. doi:10.1038/s41586-019-1799-6

11. Lång K, et al. Artificial intelligence-supported screen reading versus standard double reading in mammography screening. Lancet Oncol. 2023. Available from: https://doi.org/10.1016/S1470-2045(23)00298-X

12. Muehlematter UJ, Daniore P, Vokinger KN. Approval of artificial intelligence and machine learning-based medical devices in the USA and Europe. Lancet Digit Health. 2021. Available from: https://doi.org/10.1016/S2589-7500(20)30292-2

13. van Leeuwen KG, Schalekamp S, Rutten MJCM, van Ginneken B, de Rooij M. Artificial intelligence in radiology: 100 commercially available products and their scientific evidence. Eur Radiol. 2021. Available from: https://doi.org/10.1007/s00330-021-07892-z

14. Reichenpfader D, Muller H, Denecke K. Large language model-based information extraction from radiology reports: a scoping review. NPJ Digit Med. 2024. Available from: https://www.nature.com/npjdigitalmed/

15. Sunshine A, et al. Evaluating the quality and understandability of ChatGPT-generated radiology report summaries. JMIR Form Res. 2025. Available from: https://formative.jmir.org/2025/1/e76097/PDF

16. Agarwal N, et al. RadCliQ: clinical evaluation and quality scoring for radiology report generation. Proc Mach Learn Res. 2023. Available from: https://arxiv.org/abs/2301.06590

17. Gaur S, et al. Radiology report generation and clinical correctness: review of evaluation challenges. IEEE Rev Biomed Eng. 2024. Available from: https://doi.org/10.1109/RBME.2023.3305984

18. Asan O, Bayrak AE, Choudhury A. Artificial intelligence and human trust in healthcare: focus on clinicians. J Med Internet Res. 2020. Available from: https://doi.org/10.2196/15154

19. Glikson E, Woolley AW. Human trust in artificial intelligence: review of empirical research. Acad Manag Ann. 2020. Available from: https://doi.org/10.5465/annals.2018.0057

20. Bucinca Z, Malaya MB, Gajos KZ. To trust or to think: cognitive forcing functions can reduce overreliance on AI. Proc ACM Hum Comput Interact. 2021. Available from: https://doi.org/10.1145/3449287

21. Poursabzi-Sangdeh F, et al. Manipulating and measuring model interpretability and its impact on human-AI decision making. Proceedings of CHI. 2021. Available from: https://doi.org/10.1145/3411764.3445315

22. Tonekaboni S, Joshi S, McCradden MD, Goldenberg A. What clinicians want: contextualizing explainable machine learning for clinical end use. Machine Learning for Healthcare. 2019. Available from: https://arxiv.org/abs/1905.05134

23. Gichoya JW, et al. AI recognition of patient race in medical imaging: a modelling study. Lancet Digit Health. 2022. Available from: https://doi.org/10.1016/S2589-7500(22)00063-2

24. Seyyed-Kalantari L, et al. Underdiagnosis bias of artificial intelligence algorithms applied to chest radiographs in underserved patient populations. Nat Med. 2021. Available from: https://doi.org/10.1038/s41591-021-01595-0

25. World Health Organization. Ethics and governance of artificial intelligence for health. Geneva: World Health Organization; 2021. Available from: https://www.who.int/publications/i/item/9789240029200

26. Geis JR, et al. Ethics of artificial intelligence in radiology: summary of the joint European and North American multisociety statement. Radiology. 2019. Available from: https://doi.org/10.1148/radiol.2019191586

27. European Society of Radiology. Guiding AI in radiology: ESR's recommendations for effective application of the AI Act. Insights Imaging. 2025. Available from: https://pmc.ncbi.nlm.nih.gov/articles/PMC11825415/

28. Kelly CJ, Karthikesalingam A, Suleyman M, Corrado G, King D. Key challenges for delivering clinical impact with artificial intelligence. Lancet Digit Health. 2019. Available from: https://doi.org/10.1016/S2589-7500(19)30141-4

29. Beede E, et al. A human-centered evaluation of a deep learning system deployed in clinics for the detection of diabetic retinopathy. Proceedings of CHI. 2020. Available from: https://doi.org/10.1145/3313831.3376718

30. Creswell JW, Poth CN. Qualitative inquiry and research design: choosing among five approaches. 4th ed. Thousand Oaks: SAGE Publications; 2018. Available from: https://us.sagepub.com/en-us/nam/qualitative-inquiry-and-research-design/book246896

31. Braun V, Clarke V. Thematic analysis: a practical guide. London: SAGE Publications; 2021. Available from: https://uk.sagepub.com/en-gb/eur/thematic-analysis/book248481

32. Byrne D. A worked example of Braun and Clarke's approach to reflexive thematic analysis. Qual Quant. 2022. Available from: https://doi.org/10.1007/s11135-021-01182-y