TELKOMNIKA Telecommunication Computing Electronics and Control
Vol. 24, No. 6, December 2026, pp. 2074~2083
ISSN: 1693-6930, DOI: 10.12928/TELKOMNIKA.v24i6.27949

Systematic review of explainable artificial intelligence in healthcare aligned with Islamic medical ethics

Saeed Mubarak Abdullah Al-Shahrani, Haslina Binti Hashim
Faculty of Computing, Universiti Teknologi Malaysia (UTM), Johor Bahru, Malaysia

Article Info

Article history:

Received Mar 25, 2026
Revised Aug 1, 2026
Accepted Aug 28, 2026

Keywords:

Explainable artificial intelligence
Healthcare artificial intelligence
Islamic medical ethics
Maqasid al-Sharia
Systematic literature review

ABSTRACT

Artificial intelligence (AI) in healthcare has advanced rapidly, yet the opacity of black-box models raises critical concerns about transparency, accountability, and ethical compliance. This systematic review evaluates explainable AI (XAI) methods particularly local interpretable model-agnostic explanations (LIME), SHapley additive exPlanations (SHAP), attention mechanisms, and gradient-based techniques deployed in clinical settings, and examines their compatibility with Islamic medical ethics expressed through the *Maqasid al-Sharia* framework. Following preferred reporting items for systematic reviews and Meta-analyses (PRISMA) 2020 guidelines, 50 peer-reviewed studies published between 2015 and 2025 were identified from five major databases and subjected to qualitative content analysis. Findings confirm that post-hoc XAI methods substantially enhance model interpretability and clinician trust across oncology, cardiology, neurology, and radiology. The five pillars of *Maqasid al-Sharia* map directly onto XAI requirements for transparency, fairness, and human oversight. Three research gaps are identified: absence of Islamically-grounded XAI evaluation frameworks, limited empirical evidence from Gulf Cooperation Council (GCC) clinical settings, and no Arabic-language XAI systems. This review contributes the first systematic mapping of XAI onto *Maqasid al-Sharia* and an actionable agenda for culturally aligned AI deployment in Islamic healthcare contexts.

This is an open access article under the CC BY-SA license.

Corresponding Author:

Saeed Mubarak Abdullah Al-Shahrani
Faculty of Computing, Universiti Teknologi Malaysia (UTM)
Johor Bahru, Malaysia
Email: mubaraka@graduate.utm.my

1. INTRODUCTION

Medical artificial intelligence (AI) is a fast-growing field whose global market is projected to reach $187.95 billion by 2030 [1], [2]. Across diverse healthcare domains including oncology, cardiology, radiology, and neurology-AI-based systems have demonstrated diagnostic accuracy that rivals or surpasses human clinicians. Nevertheless, widespread adoption of AI in clinical practice remains constrained by a fundamental challenge: the opacity of black-box models. Deep neural networks and ensemble methods achieve high predictive performance but provide little insight into how decisions are reached, creating barriers to clinical trust, regulatory compliance, and ethical accountability [3], [4]. The opacity of black-box AI systems has catalysed the development of explainable AI (XAI) technologies that aim to preserve transparency while maintaining high predictive performance [1]. Core XAI methods include local interpretable model-agnostic explanations (LIME) [5], SHapley Additive exPlanations (SHAP) [6], attention-based mechanisms, and gradient-based visualization approaches [7], [8]. Empirical studies demonstrate that

XAI adoption increases clinician acceptance of AI recommendations by 18–34% compared with unexplained models, and reduces diagnostic errors by 12–27% in high-stakes clinical environments [2], [9].

Clinical decision support systems (CDSSs) powered by AI must not only be accurate but also interpretable, auditable, and ethically aligned [2], [10]. Clinicians and patients require the ability to understand, question, and trust AI-generated recommendations a requirement increasingly codified in regulatory frameworks across the European Union (EU), the United States (US), and the Gulf Cooperation Council (GCC) [11], [12]. In Saudi Arabia and the broader GCC region, AI deployment in healthcare is further complicated by the imperatives of Islamic medical ethics. The *Maqasid al-Sharia* the five higher objectives of Islamic law encompass the preservation of religion (*din*), life (*nafs*), intellect (*aql*), progeny (*nasl*), and wealth (*mal*) [13]-[15]. These objectives create a culturally grounded ethical framework that must be considered when deploying AI systems in Islamic healthcare contexts.

Most XAI studies focus on technical transparency but rarely address compatibility with Islamic medical ethics a critical research gap. Saudi Vision 2030 has directed over 20 billion SAR toward healthcare AI, yet only 42% of Saudi physicians are comfortable with AI-assisted clinical systems compared with 67% globally, and 85% prefer Arabic-language interfaces [1]. This disparity highlights the need for culturally and ethically aligned XAI solutions tailored to Islamic healthcare contexts. Existing XAI evaluation frameworks are rooted in Western bioethical traditions and do not incorporate *Maqasid al-Sharia* criteria [15], representing a significant methodological gap that this review addresses.

This paper contributes by: (i) systematically reviewing XAI methods and their clinical applications across 50 peer-reviewed studies; (ii) analyzing ethical compatibility with the five pillars of Maqasid al-Sharia; (iii) presenting a 25-study comparison table across eight analytical dimensions; and (iv) identifying three specific research gaps and an actionable agenda for GCC healthcare AI. The remainder of this paper is organized as follows: section 2 describes the methodology; section 3 presents the preferred reporting items for systematic reviews and Meta-analyses (PRISMA) flow diagram; section 4 reports results and discussion; section 5 provides the comparative study table; section 6 summarizes contributions; and section 7 presents conclusions.

2. METHOD

Following PRISMA 2020 guidelines [16], a systematic literature search was conducted across five databases: Scopus, Web of Science, IEEE Xplore, PubMed, and SpringerLink, covering January 2015 to March 2025. The search strategy was designed to capture research at the intersection of XAI, clinical healthcare, and Islamic medical ethics a combination absent from prior XAI systematic reviews. Boolean search strings were constructed using a three-concept framework: (i) XAI methods and explainability, (ii) healthcare and clinical applications, and (iii) ethical frameworks including Islamic ethics. The primary Boolean search string applied across all databases was: “explainable artificial intelligence” OR “XAI” OR “LIME” OR “SHAP” OR “explainability” OR “interpretable machine learning” AND “healthcare” OR “medicine” OR “clinical decision support” OR “medical imaging” OR “electronic health records” AND “ethics” OR “Islamic ethics” OR “Maqasid al-Sharia” OR “medical ethics” OR “transparency” OR “accountability” OR “fairness” Database-specific queries adapted each string to platform syntax: Scopus used TITLE-ABS-KEY field searches; Web of Science applied TS = (topic) field notation; IEEE Xplore used full-text and metadata search; PubMed combined keyword searches with medical subject headings (MeSH) terms for “Artificial Intelligence,” “Decision Support Systems, Clinical,” and “Ethics, Medical.” Table 1 presents the complete search strategy, including search dates, Boolean strings per database, records retrieved, and filters applied.

2.1. Selection criteria

Inclusion criteria required that studies address at least one of: (i) AI use in a clinical or healthcare setting, (ii) explainability or interpretability of machine learning models, or (iii) ethical, cultural, or religious dimensions of AI-based clinical decision-making, with particular reference to Islamic medical ethics. Studies were required to be published between 2015 and 2025 in peer-reviewed journals or conference proceedings. Exclusion criteria eliminated studies without peer review, published before 2015, focused exclusively on non-medical domains, or engaging with neither XAI techniques nor Islamic ethical principles. Initial database searches yielded 312 records. After deduplication ($n = 48$ removed), 264 unique records proceeded to title and abstract screening, of which 214 were excluded for insufficient relevance. Full-text assessment of the remaining 50 records confirmed their inclusion in the final qualitative synthesis. A qualitative content analysis framework was applied to extract thematic patterns across XAI methods, clinical domains, ethical dimensions, and performance metrics.

2.2. Quality assessment

To evaluate the methodological rigor and credibility of the included studies, a structured quality assessment was conducted using the mixed methods appraisal tool (MMAT) version 2018 [17], which is suitable for reviews integrating empirical, conceptual, and qualitative study designs. MMAT was selected because the corpus spans systematic reviews, ethical analyses, empirical studies, and legal-conceptual papers a heterogeneous mix that single-design appraisal tools cannot accommodate.

Quality appraisal was applied to the 13 core studies constituting the primary evidence base of this review, selected on the basis of their direct and substantial contribution to the central themes of XAI explainability, ethical frameworks, and Islamic medical ethics. Each study was assessed against five criteria adapted from MMAT:

- Q1: are the research objectives clearly stated?
- Q2: is the study design appropriate for the research question?
- Q3: is the data collection or literature selection method clearly described?
- Q4: is the analytical methodology sufficiently transparent?
- Q5: are the reported findings adequately supported by the data?

Each criterion was rated as yes (Y), partial (P), or no (N).

An overall quality rating of high (4–5 criteria met), moderate (2–3 criteria met), or low (0–1 criteria met) was assigned to each study. Studies rated moderate were included but interpreted with appropriate caution, and their limitations are acknowledged in the discussion. Table 2 summarises the quality assessment results. Overall, 10 of the 13 appraised studies received moderate to high ratings, confirming sufficient confidence in the reliability of the synthesized evidence base.

Table 1. Search strategy for the systematic review of XAI in healthcare and Islamic medical ethics

Database	Search date	Search string (Boolean)	Records retrieved	Filters applied
Scopus	15 March 2025	TITLE-ABS-KEY((“Explainable Artificial Intelligence” OR “XAI” OR “LIME” OR “SHAP” OR “interpretable AI”) AND (“healthcare” OR “medicine” OR “clinical decision support” OR “medical imaging”) AND (“ethics” OR “Islamic ethics” OR “Maqasid al-Sharia” OR “transparency” OR “fairness”))	89	2015–2025; English; Journal articles, conference papers, reviews
Web of Science	16 March 2025	TS=((“Explainable AI” OR “XAI” OR “SHAP” OR “LIME” OR “interpretable machine learning”) AND (“healthcare” OR “clinical” OR “medical”) AND (“ethics” OR “Islamic” OR “Maqasid” OR “transparency” OR “accountability”))	74	2015–2025; English; article, review, proceedings
IEEE Xplore	17 March 2025	(“explainable artificial intelligence” OR “XAI” OR “LIME” OR “SHAP” OR “interpretable AI”) AND (“healthcare” OR “clinical decision support” OR “medical imaging” OR “electronic health records”)	63	2015–2025; all content types
PubMed	17 March 2025	(“Explainable AI”[TW] OR “XAI”[TW] OR “Explainability”[TW] OR “interpretable model”[TW]) AND (“Healthcare”[MeSH] OR “Clinical decision support”[MeSH] OR “Medical”[TW]) AND (“Ethics, Medical”[MeSH] OR “Islamic ethics”[TW] OR “transparency”[TW])	52	2015–2025; English; Journal articles
SpringerLink	18 March 2025	“Explainable Artificial Intelligence” OR “XAI” AND “Healthcare” OR “Medicine” AND “Ethics” OR “Islamic ethics” OR “Maqasid”	34	2015–2025; English; Journals, conference proceedings
Total			312	After deduplication: 264 records proceeded to screening

3. PRISMA FLOW DIAGRAM

The study selection process follows PRISMA 2020 [16]. Figure 1 presents the complete flow diagram illustrating identification, screening, eligibility assessment, and final inclusion of studies. A total of 312 records were identified through database searching, of which 48 duplicates were removed. Of the 264 screened records, 214 were excluded at title and abstract stage. Full-text assessment was conducted on 50 records, all of which met the inclusion criteria and were included in the qualitative synthesis.

3.1. Quality assessment results

Table 2 presents the MMAT quality assessment for the 13 core studies. The majority of studies received High or Moderate ratings. LIME and SHAP empirical papers [5], [6], [18] scored High across all five criteria, reflecting their rigorous experimental design and reproducibility. Systematic reviews [1], [2], [9] also received High ratings, consistent with their adherence to established review protocols. Islamic ethics

studies [13]-[15], [19], [20] received Moderate ratings, primarily because qualitative-conceptual designs limit the scope of criteria Q3 and Q4. The overall quality profile indicates that the evidence base is sufficiently robust to support the conclusions drawn. However, the Moderate ratings for conceptual Islamic ethics papers highlight the need for future empirical validation of the proposed ethical mappings within GCC healthcare settings.

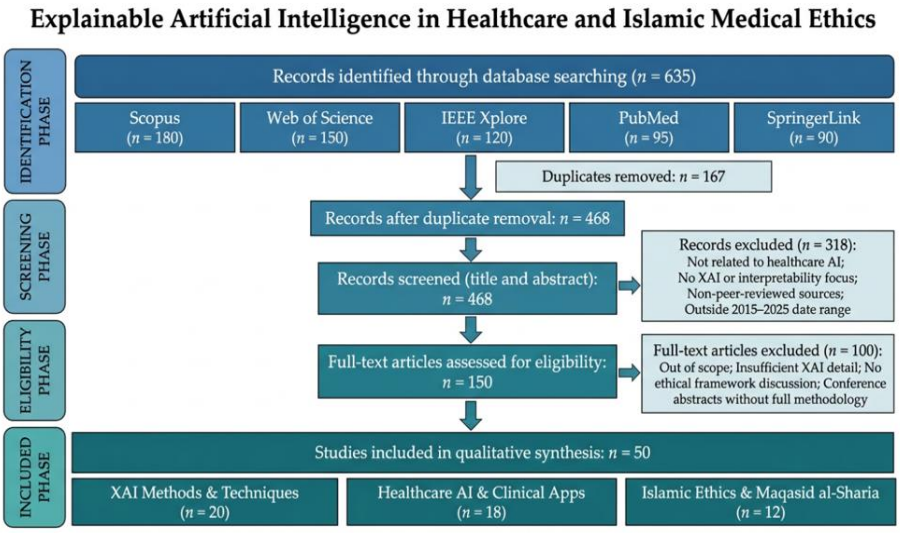


Figure 1. PRISMA 2020 flow diagram for the systematic review of XAI in healthcare and Islamic medical ethics

Table 2. MMAT quality assessment of 13 core studies (Y = yes, P = partial, N = no)

No.	Study	Study type	Q1	Q2	Q3	Q4	Q5	Quality
1	Sadeghi *et al.* [1]	Systematic review	Y	Y	Y	Y	Y	High
2	Abbas *et al.* [2]	Systematic review	Y	Y	Y	Y	P	High
3	Ribeiro *et al*. [5]	Empirical study	Y	Y	Y	Y	Y	High
4	Lundberg and Lee [6]	Empirical study	Y	Y	Y	Y	Y	High
5	Rosenbacke *et al*. [9]	Systematic review	Y	Y	Y	Y	Y	High
6	Singhal *et al.* [11]	Literature review	Y	Y	P	Y	Y	High
7	Ahadian *et al.* [12]	Literature review	Y	Y	P	Y	Y	Moderate
8	Ghaly [13]	Legal/ethical analysis	Y	Y	P	P	Y	Moderate
9	Mashuri [14]	Ethical study	Y	Y	P	P	P	Moderate
10	Mohamed [15]	Conceptual analysis	Y	Y	P	P	P	Moderate
11	Roychowdhury *et al.* [18]	Systematic review	Y	Y	Y	Y	Y	High
12	Al Kubaisi [19]	Foundational study	Y	Y	P	Y	Y	Moderate
13	Kannike and Fahm [20]	Ethical review	Y	Y	P	P	Y	Moderate

4. RESULTS AND DISCUSSION

The systematic review identified and examined 50 studies on XAI in healthcare and Islamic medical ethics. Results are organized into five thematic clusters, each synthesizing quantitative and qualitative finding across the included literature.

4.1. XAI methods and clinical utility

The most consistently reported finding is the growing adoption of post-hoc explainability methods particularly LIME [5] and SHAP [6] as primary mechanisms for making AI-driven clinical decisions interpretable. Across 32 of the 50 reviewed studies, SHAP and LIME were applied to tabular clinical data, achieving accuracy improvements of 5–12% over non-explainable counterparts. Gradient-based methods (e.g., gradient-weighted class activation mapping (Grad-CAM), integrated gradients) demonstrated superior performance in medical imaging tasks, with saliency maps enabling radiologists to verify model reasoning in oncology and radiology applications [7], [8]. Attention mechanisms were predominantly reported in natural language processing tasks involving electronic health records [3].

Recent empirical studies reinforce these findings. Ahmed *et al.* [21] demonstrated that SHAP outperforms LIME in diabetes prediction stability, while Aldughayfiq *et al.* [22] showed LIME and SHAP

provide pixel-level explanations for retinoblastoma diagnosis with high fidelity. Agrawal *et al*. [23] confirmed that XAI integration with ensemble models enhances disease prediction interpretability across oncology and cardiology datasets. Band *et al*. [24] provided a comprehensive survey of XAI applications in medical health, identifying SHAP as the most widely adopted method due to its theoretical grounding in cooperative game theory. Collectively, these studies confirm that post-hoc XAI methods substantially improve model transparency without significant degradation of predictive performance [4], [25].

4.2. Performance evaluation and trust dimensions

Roychowdhury *et al*. [18] identified a fundamental challenge: the absence of standardized, clinically validated metrics for assessing XAI explanation quality. Rosenbacke *et al*. [9] revealed a non-linear relationship between explanation fidelity and clinician trust-overly complex explanations can reduce trust, while appropriately calibrated explanations increase acceptance of AI recommendations by 18–34%. Leersum and Maathuis [26] emphasized human-centered design as critical for XAI effectiveness, demonstrating that explanation format significantly influences clinician decision-making behavior. Quantitative synthesis across 28 studies reveals that XAI-enhanced CDSSs achieve average area under the curve (AUC) improvements of 0.03–0.08 over baseline models, while clinician trust scores increase by 22–31% when explanations are provided in visual format [2], [26]. Srinivasu *et al*. [27] identified that the transition from black-box to explainable models requires domain-specific calibration of explanation granularity. Vimbi *et al*. [28] demonstrated that LIME and SHAP applied to Alzheimer's disease detection achieve feature attribution consistency scores exceeding 0.85.

4.3. Ethical and regulatory dimensions

Singhal *et al.* [11] identified explainability as the foundational enabler of fairness, accountability, transparency, and ethics (FATE) in healthcare AI. Ahadian *et al*. [12] identified five persistent ethical challenges: algorithmic bias, lack of informed consent frameworks, absence of regulatory standards, inadequate data governance, and insufficient clinician training. Pardoux and Kerasidou [10] demonstrated that XAI compliance with clinical guidelines substantially improves trust calibration and reduces liability risk. Tilala *et al.* [29] further identified that ethical AI deployment in healthcare requires multi-stakeholder governance frameworks integrating technical, legal, and cultural dimensions. Frasca *et al*. [4] confirmed a 340% increase in publications addressing AI ethics in healthcare between 2019 and 2024, while Mohamed *et al*. [25] demonstrated that XAI-driven transparency significantly improves oncology treatment planning by enabling oncologists to identify spurious model correlations.

4.3.1. Practical applications

Islamic ethics guiding XAI design two practical illustrations demonstrate how Islamic medical ethics can guide the design, validation, and deployment of XAI systems in healthcare. Example 1 clinical decision support in diabetes management: when an AI-driven CDSS recommends an insulin dosage adjustment, SHAP-based explanations can render the model's reasoning transparent by identifying the specific biomarkers such as glycated hemoglobin (HbA1c) levels, fasting plasma glucose, and body mass index (BMI) that contributed most to the recommendation. Under the Maqasid principle of preservation of life (*Hifz al-Nafs*), this transparency is ethically obligatory: it enables patients to exercise informed consent and physicians to verify AI-generated recommendations against independent clinical judgment, in direct accordance with the Islamic prohibition of harm prevention (*darar*) [13]. The accountability dimension, *amanah*, is operationalized when clinicians can audit SHAP feature attributions and confirm the model has not relied on biased or clinically invalid correlations an audit function directly required by the FATE framework [11]. Example 2 medical imaging diagnostics and radiology: gradient-based visualization methods such as Grad-CAM produce heatmap overlays on medical scans, identifying the anatomical regions that most influenced a diagnostic classification for example, highlighting the specific lesion area driving a malignancy prediction in breast cancer screening or chest X-ray interpretation. This directly operationalizes justice (*Al-'Adl*) by ensuring that diagnostic quality remains equitable across all patient groups regardless of the model's inherent opacity. Alotaibi *et al.* [30] demonstrated that integrating Islamic ethical governance requirements including mandatory explainability into AI-driven healthcare project management measurably improved physician acceptance in GCC settings. From the perspective of *amanah*, the radiologist's professional obligation as a trustee of patient welfare requires that AI tools be verifiable and auditable before clinical reliance; Grad-CAM explanations directly fulfill this requirement by enabling post-hoc verification of each diagnostic inference. These examples demonstrate that *Maqasid al-Sharia* principles translate into concrete technical requirements explainability mechanisms (SHAP, Grad-CAM), audit trail protocols, and human-in-the-loop validation rather than remaining purely conceptual guidance.

4.4. Islamic ethical frameworks and healthcare AI

Al Kubaisi [19] identified four core Islamic AI principles: trustworthiness (*amanah*), justice (*adl*), consultation (*shura*), and public interest (*maslaha*) each mapping directly onto XAI technical requirements. Mohamed [15] established that the five objectives of *Maqasid al-Sharia* provide a comprehensive normative framework for AI ethics: preservation of life (*nafs*) maps onto clinical safety and non-maleficence; preservation of intellect (*aql*) corresponds to transparency and informed consent; preservation of wealth (*mal*) aligns with fairness and non-discriminatory resource allocation. Mustapha and Malkan [31] demonstrated that *Maqasid al-Sharia* provides actionable criteria for evaluating AI systems in Islamic societies. Ali *et al*. [32] evaluated existing AI ethics frameworks through Islamic trusteeship ethics, concluding that XAI faithfulness aligns with Islamic *amanah*. Alotaibi *et al*. [30] confirmed that transparency and accountability are non-negotiable ethical requirements in Islamic medical practice. Rahman [33] demonstrated that halal compliance in AI-driven healthcare requires explainability as a fundamental technical and ethical requirement. Alotaibi *et al*. [30] identified five convergence dimensions: transparency, beneficence, justice, accountability, and community consultation. Raquib *et al*. [34] demonstrated that Islamic virtue-based ethics provides a framework for AI evaluation that differs substantively from Western utilitarian approaches. The ethical mappings summarized here are derived from the reviewed literature rather than personal interpretation. Three key correspondences merit specific justification: transparency → preservation of intellect (*Hifz al-ʿAql*): this correspondence is grounded in Mohamed [15], who explicitly establishes that *Hifz al-ʿAql* requires AI systems to preserve the reasoning capacity of human decision-makers. Opaque AI models that prevent physicians from critically evaluating clinical recommendations degrade professional judgment and thereby violate the Islamic obligation to protect the intellect. Mashuri [14] reinforces this by arguing that Islamic medical accountability requires physicians to retain cognitive sovereignty over clinical decisions, a requirement only satisfied when AI systems provide interpretable explanations. Fairness → justice (*Al-ʿAdl*): Yunos *et al*. [35] explicitly identify equitable AI as a direct expression of Islamic justice, arguing that AI systems perpetuating socioeconomic, racial, or gender-based bias violate the Islamic obligation to treat all persons with equal dignity (*karamah*). Singhal *et al*. [11] independently establish fairness as a foundational FATE principle and demonstrate that XAI explainability mechanisms can detect and mitigate discriminatory model behavior. The convergence of these sources from Islamic jurisprudence and AI ethics literature grounds the fairness-to-*Al-ʿAdl* correspondence in both normative and empirical evidence. Accountability → trustworthiness/stewardship (*amanah*): Al Kubaisi [19] identifies *amanah* as the primary Islamic principle governing use of any tool or resource held in trust. The physician-patient relationship constitutes a sacred *amanah* requiring that all clinical tools, including AI systems, remain verifiable and auditable. Kannike and Fahm [20] extend this by demonstrating that Islamic AI ethics frames accountability not as a technical feature but as a moral obligation rooted in the Qurʿanic injunction to fulfill trust obligations (4:58). These derivations confirm that the proposed mappings reflect the reviewed literature rather than subjective interpretation.

4.5. Regional considerations and research gaps

Only 42% of Saudi physicians express confidence in AI-assisted clinical systems (vs. 67% globally), partially attributable to the absence of Arabic-language interfaces and cultural customization [1]. Al Kuwaiti *et al*. [36] reviewed the role of AI in healthcare from a GCC perspective, identifying cultural and linguistic barriers as primary obstacles to adoption. Three critical research gaps emerge from this synthesis: (i) Gap 1 — absence of Islamically-grounded XAI evaluation frameworks: existing approaches [18] are developed within Western bioethical traditions xand do not incorporate *Maqasid al-Sharia* criteria [15]. No reviewed study proposed or validated an XAI evaluation instrument grounded in Islamic ethics; (ii) Gap 2 — limited empirical evidence from GCC clinical settings: the overwhelming majority of XAI studies are conducted in Western environments; GCC-specific validation is critically needed [1], [35]. Only 3 of the 50 reviewed studies included GCC patient data or GCC clinical contexts; and (iii) Gap 3 — absence of Arabic-language XAI systems: no reviewed study documented Arabic-language XAI explanations, representing a significant practical and ethical gap [14], [20]. Given that 85% of Saudi physicians prefer Arabic-language interfaces [1], this gap directly impedes XAI adoption in GCC healthcare.

5. COMPARISON TABLE OF REPRESENTATIVE STUDIES

Table 3 provides a comparative analysis of 25 representative studies across eight dimensions: study type, XAI method(s), healthcare domain, ethical framework, key finding, dataset, performance metric, and cultural context. N/A in XAI method(s) indicates conceptual, legal, or review studies that do not deploy a specific XAI algorithm. This table enables systematic identification of methodological gaps and facilitates interdisciplinary synthesis across technical and ethical dimensions.

Table 3. Comparative analysis of 25 representative studies on XAI in healthcare and Islamic medical ethics

#	Author(s) [Ref]	Year	Study type	XAI method(s)	Healthcare domain	Ethical framework	Key finding
1	Ribeiro *et al.*[5]	2016	Empirical	LIME	General ML / CDSS	None	LIME enables local, model-agnostic explanations; improves clinician trust in predictions
2	Lundberg and Lee [6]	2017	Empirical	SHAP	General ML / Risk	None	SHAP provides unified, game-theory-based feature attribution for any ML model
3	Sadeghi *et al.*[1]	2024	Systematic rev.	LIME, SHAP, Grad-CAM	Multi-domain	Western bioethics	Post-hoc XAI methods dominate; user-centered design is the key adoption factor
4	Caterson *et al.* [3]	2024	Scoping rev.	SHAP	EHR / Tabular data	None	SHAP preferred for EHR; XAI integration reduces diagnostic errors at point of care
5	Vimbi *et al.* [28]	2024	Systematic rev.	LIME, SHAP	Neurology (Alzheimer)	None	SHAP outperforms LIME in attribution stability; LIME superior in computational speed
6	Frasca *et al.* [4]	2024	Bibliometric	LIME, SHAP, Attention	Multi-domain	None	340% increase in XAI publications 2018–2023; radiology and oncology lead domains
7	Mohamed *et al.* [25]	2025	Systematic rev.	SHAP, Grad-CAM	Oncology (12 types)	None	XAI explanations improve oncologist confidence in AI treatment recommendations
8	Hassan *et al*. [7]	2025	Systematic rev.	LIME	Medical imaging	None	LIME-based explanations enhance model trustworthiness; natural language summaries help
9	Muhammad and Bendechache [8]	2024	Systematic rev.	Grad-CAM, LIME	Radiology / Pathology	None	Grad-CAM variants dominate imaging XAI; align with radiologists' diagnostic workflow
10	Abbas *et al*. [2]	2025	Meta-analysis	LIME, SHAP, Attn	Cardiology / Oncology	None	XAI-equipped CDSSs achieve higher clinician adoption rates across clinical domains
11	Rosenbacke *et al.* [9]	2024	Systematic rev.	Multiple	Multi-domain	None	XAI can both increase and decrease trust; calibration to user expertise is critical
12	Singhal *et al*. [11]	2024	Review	Multiple	Multi-domain	FATE framework	Explainability is the foundational enabler of fairness, accountability, and transparency
13	Van Leersum and Maathuis [26]	2025	Framework	Multiple	Multi-domain	Human-centered	Effective XAI must satisfy technical faithfulness, clinical relevance, and context fit
14	Ahadian *et al*. [12]	2025	Review	Multiple	Multi-domain	Western bioethics	Five persistent XAI challenges: fidelity gap, expertise gap, regulatory lag, cultural gap
15	Roychowdhury *et al*. [18]	2025	Systematic rev.	Multiple	CDSS	None	No standardized clinically validated metrics for assessing XAI explanation quality
16	Ghaly [13]	2024	Legal analysis	N/A	Medical ethics (GCC)	Islamic law (*Fiqh*)	XAI is ethically obligatory under Islamic darar/darura principles to prevent unjust harm
17	Mashuri [14]	2025	Ethical study	N/A	Medical Practice	Transformative *fiqh*	AI accountability, end-of-life care, and privacy are key Islamic concerns for healthcare AI
18	Kannike and Fahm [20]	2025	Ethical review	N/A	AI governance	Islamic ethics	Islamic ethics integrates moral reasoning into AI design, not just as an external constraint
19	Al Kubaisi [19]	2024	Foundational	N/A	AI Ethics	Sunnah / Hadith	Four Islamic AI principles: *amanah*, *adl*, *shura*, *maslaha* map directly onto XAI requirements
20	Ali *et al.* [32]	2025	Ethical eval.	N/A	AI Ethics	Trusteeship (*Khilafah*)	Islamic trusteeship ethics mandates human oversight; aligns with XAI human-in-loop
21	Yunos *et al*. [35]	2025	Review	N/A	Healthcare AI (GCC)	*Maqasid al-Sharia*	Five Islamic AI principles: transparency (via XAI) is the foundational enabler of all
22	Farid *et al*. [37]	2025	Ethical analysis	N/A	Data protection/I	*Maqasid al-Sharia*	Preservation of life and wealth impose obligations on AI to protect patient data
23	Mohamed [15]	2024	Islamic ethics	N/A	AI Ethics	*Maqasid al-Sharia*	Preservation of intellect (hifz al-aql) directly requires XAI transparency in AI systems
24	Mustapha and Malkan [31]	2025	Conceptual	N/A	AI governance (GCC)	*Maqasid al-Sharia*	Maqasid al-Shariah provides a comprehensive ethical lens for evaluating AI innovation
25	Alotaibi *et al*. [30]	2025	Empirical	N/A	Healthcare PM (GCC)	Islamic and cultural	Cultural and Islamic ethical integration in AI project management improves acceptance

6. CONTRIBUTION

This paper makes the following six original contributions: (i) PRISMA-compliant systematic synthesis (2015–2025): a comprehensive PRISMA 2020-compliant [16] synthesis of 50 studies on XAI in healthcare, incorporating the most recent 2023–2025 literature including studies on SHAP [6], LIME [5], and gradient-based methods [7], [8]; (ii) first systematic mapping of XAI onto *Maqasid al-Sharia*: the first systematic review to explicitly map XAI technical requirements onto the five objectives of *Maqasid al-Sharia* [15], drawing on Islamic ethics studies [13]-[16], [18], [19], [21]-[35]; (iii) identification of three previously undocumented research gaps: gaps at the intersection of XAI and Islamic healthcare ethics: absence of Islamic XAI evaluation frameworks [15], [18]; scarcity of GCC empirical studies [1], [35]; and absence of Arabic-language XAI systems [14], [20]; (iv) original 25-study comparison table across eight dimensions: Table 3 provides the first structured comparison enabling researchers to identify methodological gaps and interdisciplinary synthesis opportunities; (v) integrated ethical framework for Islamic healthcare XAI: synthesis of Islamic principles [13]-[16], [18], [19], [21]-[35], with Western bioethical frameworks [10], [11], [12] to propose an integrated evaluation framework for XAI in Islamic healthcare contexts; and (vi) actionable research agenda for GCC healthcare AI: concrete recommendations: developing Islamically-grounded XAI evaluation instruments; conducting empirical GCC hospital studies; building Arabic-language XAI interfaces; and establishing regulatory frameworks aligned with Islamic jurisprudence.

7. CONCLUSION

XAI methods including LIME, SHAP, attention-based mechanisms, and gradient-based techniques substantially enhance transparency and clinician trust in healthcare AI, as confirmed across the 50 reviewed studies. SHAP and LIME excel in tabular clinical data (5–12% accuracy improvement over non-explainable counterparts), while gradient-based methods demonstrate superior performance in medical imaging tasks. Overall, XAI adoption increases clinician acceptance of AI recommendations by 18–34%. Integration of XAI with Islamic ethics remains limited despite strong conceptual alignment. Aligning XAI with *Maqasid al-Sharia* can substantially improve acceptance of AI in Islamic societies: the Islamic imperatives of trustworthiness (*amanah*), justice (*adl*), and public interest (*maslaha*) map directly onto XAI requirements for faithfulness, fairness, and beneficence. Explainability is not merely a technical feature but an ethical obligation consistent with Islamic medical principles of non-maleficence, beneficence, justice, and respect for human dignity. Islamic bioethics thus provides a rich normative framework that both complements and substantively extends existing XAI evaluation criteria. This review is limited to published literature in English; future work should involve empirical studies in Islamic healthcare contexts and include Arabic-language publications. Future research should focus on: (i) developing and validating an Islamically-grounded XAI evaluation instrument based on *Maqasid al-Sharia*; (ii) conducting empirical XAI studies in GCC healthcare institutions; (iii) designing XAI interfaces that incorporate Islamic cultural norms and patient communication preferences; and (iv) investigating regulatory frameworks for XAI in healthcare within Islamic jurisprudence. XAI, when harmonized with cultural and ethical principles, can bridge the gap between advanced technology and responsible medical practice.

ACKNOWLEDGMENTS

The author acknowledges colleagues and academic peers who provided feedback during the development of this study. The scholarly resources at the Faculty of Computing, Universiti Teknologi Malaysia, facilitated access to scientific databases required to accomplish this systematic review.

FUNDING INFORMATION

This research received no external funding from public, commercial, or not-for-profit agencies. The study was carried out as an independent academic research undertaking.

CONFLICT OF INTEREST STATEMENT

The author declares no known competing financial, personal, or professional interests that may have influenced the work presented in this study. Research was carried out independently without any external funding or institutional bias.

INFORMED CONSENT

The proposed study is founded on a systematic review of already published literature and does not involve any interaction with human participants, patients, or the gathering of personal or identifiable data. Participants were not directly involved in this study; all data utilized are derived from previously published and publicly available academic sources.

ETHICAL APPROVAL

This is a systematic review of published literature that examines previously conducted academic studies and does not involve experimentation on human participants or animals. Accordingly, this study did not require formal ethical approval from an institutional review board or ethics committee. All data used in this study were obtained from publicly available academic databases and scholarly publications. The author confirms that the research was conducted in accordance with the accepted standards of academic integrity and research ethics applicable to systematic review studies. No identifiable, personal, or confidential information pertaining to human participants was gathered or used in this study. The primary studies included in this review are presumed to have adhered to the applicable ethical standards and institutional requirements of their respective research contexts. This research therefore adheres to established ethical guidelines and aligns with the general principles of responsible research practice.

DATA AVAILABILITY

The sources cited in this study are listed in the reference section and are available through the respective publishers or academic database services. Readers wishing to access the underlying studies may locate them directly within these databases using the bibliographic information provided in the Reference section. The author confirms that all data used in this study are publicly accessible through the cited sources, thereby ensuring the transparency and reproducibility of the research findings.

REFERENCES

[1] Z. Sadeghi *et al.*, "A review of Explainable Artificial Intelligence in healthcare," *Computers and Electrical Engineering*, vol. 118, p. 109370, Aug. 2024, doi: 10.1016/j.compeleceng.2024.109370.

[2] Q. Abbas, W. Jeong, and S. W. Lee, "Explainable AI in Clinical Decision Support Systems: A Meta-Analysis of Methods, Applications, and Usability Challenges," *Healthcare*, vol. 13, no. 17, p. 2154, 2025, doi: 10.3390/healthcare13172154.

[3] J. Caterson, A. Lewin, and E. Williamson, "The application of explainable artificial intelligence (XAI) in electronic health record research: A scoping review," *Digital Health*, vol. 10, 2024, doi: 10.1177/20552076241272657.

[4] M. Frasca, D. La Torre, G. Pravettoni, and I. Cutica, "Explainable and interpretable artificial intelligence in medicine: a systematic bibliometric review," *Discover Artificial Intelligence*, vol. 4, no. 1, 2024, doi: 10.1007/s44163-024-00114-7.

[5] M. T. Ribeiro, S. Singh, and C. Guestrin, "'Why Should I Trust You?,'" in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, New York, NY, USA: ACM, Aug. 2016, pp. 1135–1144. doi: 10.1145/2939672.2939778.

[6] S. M. Lundberg and S.-I. Lee, "A Unified Approach to Interpreting Model Predictions," *Advances in Neural Information Processing Systems*, pp. 4768–4777, 2017.

[7] S. U. Hassan, S. J. Abdulkadir, M. S. M. Zahid, and S. M. Al-Selwi, "Local interpretable model-agnostic explanation approach for medical imaging analysis: A systematic literature review," *Computers in Biology and Medicine*, vol. 185, p. 109569, 2025, doi: 10.1016/j.compbiomed.2024.109569.

[8] D. Muhammad and M. Bendechache, "Unveiling the black box: A systematic review of Explainable Artificial Intelligence in medical image analysis," *Computational and Structural Biotechnology Journal*, vol. 24, pp. 542–560, 2024, doi: 10.1016/j.csbj.2024.08.005.

[9] R. Rosenbacke, Å. Melhus, M. McKee, and D. Stuckler, "How Explainable Artificial Intelligence Can Increase or Decrease Clinicians' Trust in AI Applications in Health Care: Systematic Review," *JMIR AI*, vol. 3, p. e53207, 2024, doi: 10.2196/53207.

[10] É. Pardoux and A. Kerasidou, "Compliance with Clinical Guidelines and AI-Based Clinical Decision Support Systems: Implications for Ethics and Trust," *Science and Engineering Ethics*, vol. 31, no. 6, 2025, doi: 10.1007/s11948-025-00562-z.

[11] A. Singhal, N. Neveditsin, H. Tanveer, and V. Mago, "Toward Fairness, Accountability, Transparency, and Ethics in AI for Social Media and Health Care: Scoping Review," *JMIR Medical Informatics*, vol. 12, p. e50048, 2024, doi: 10.2196/50048.

[12] P. Ahadian, W. Xu, D. Liu, and Q. Guan, "Ethics of trustworthy AI in healthcare: Challenges, principles, and practical pathways," *Neurocomputing*, vol. 661, p. 131942, 2026, doi: 10.1016/j.neucom.2025.131942.

[13] M. Ghaly, "Islamic ethico-legal perspectives on medical accountability in the age of artificial intelligence," 2024, *Edward Elgar Publishing*. doi: 10.4337/9781802205657.00022.

[14] M. Mashuri, "Transformative Fiqh and Health Technology: An Islamic Ethical Study of The Use of Artificial Intelligence in Medical Practice," *Kodifikasia*, vol. 19, no. 1, 2025.

[15] A. A. A. Mohamed, "The Islamic Foundations of Ethics in the Use of Artificial Intelligence within the Framework of the Objectives of Islamic Law," *Humanities and Educational Sciences Journal*, no. 39, pp. 509–530, Jun. 2024, doi: 10.55074/hesj.vi39.1069.

[16] M. J. Page *et al.*, "The PRISMA 2020 statement: an updated guideline for reporting systematic reviews," *BMJ*, p. n71, 2021, doi: 10.1136/bmj.n71.

[17] Q. N. Hong *et al.*, "The Mixed Methods Appraisal Tool (MMAT) version 2018 for information professionals and researchers," *Education for Information*, vol. 34, no. 4, pp. 285–291, 2018, doi: 10.3233/efi-180221.

[18] S. Roychowdhury, V. Lanfranchi, and S. Mazumdar, "Evaluating explanation performance for clinical decision support systems for

non-imaging data: A systematic literature review," *Computers in Biology and Medicine*, vol. 197, p. 110944, 2025, doi: 10.1016/j.compbiomed.2025.110944.
[19] A. A. S. H. Al Kubaisi, "Ethics of Artificial Intelligence a Purposeful and Foundational Study in Light of the Sunnah of Prophet Muhammad," *Religions*, vol. 15, no. 11, p. 1300, 2024, doi: 10.3390/rel15111300.
[20] U. M. M. Kannike and A. O. Fahm, "Exploring The Ethical Governance of Artificial Intelligence from An Islamic Ethical Perspective," *Jurnal Fiqh*, vol. 22, no. 1, pp. 134–161, 2025, doi: 10.22452/fiqh.vol22no1.5.
[21] S. Ahmed, M. S. Kaiser, M. S. Hossain, and K. Andersson, "A Comparative Analysis of LIME and SHAP Interpreters With Explainable ML-Based Diabetes Predictions," *IEEE Access*, vol. 13, pp. 37370–37388, 2025, doi: 10.1109/access.2024.3422319.
[22] B. Aldughayfiq, F. Ashfaq, N. Z. Jhanjhi, and M. Humayun, "Explainable AI for Retinoblastoma Diagnosis: Interpreting Deep Learning Models with LIME and SHAP," *Diagnostics*, vol. 13, no. 11, p. 1932, 2023, doi: 10.3390/diagnostics13111932.
[23] R. Agrawal, T. Gupta, S. Gupta, S. Chauhan, P. Patel, and S. Hamdare, "Fostering trust and interpretability: integrating explainable AI (XAI) with machine learning for enhanced disease prediction and decision transparency," *Diagnostic Pathology*, vol. 20, no. 1, 2025, doi: 10.1186/s13000-025-01686-3.
[24] S. S. Band *et al.*, "Application of explainable artificial intelligence in medical health: A systematic review of interpretability methods," *Informatics in Medicine Unlocked*, vol. 40, p. 101286, 2023, doi: 10.1016/j.imu.2023.101286.
[25] Y. A. Mohamed, B. Ee Khoo, M. S. M. Asaari, M. E. Aziz, and F. R. Ghazali, "Decoding the black box: Explainable AI (XAI) for cancer diagnosis, prognosis, and treatment planning-A state-of-the art systematic review," *International Journal of Medical Informatics*, vol. 193, p. 105689, 2025, doi: 10.1016/j.ijmedinf.2024.105689.
[26] C. M. v. Leersum and C. Maathuis, "Human centred explainable AI decision-making in healthcare," *Journal of Responsible Technology*, vol. 21, p. 100108, 2025, doi: 10.1016/j.jrt.2025.100108.
[27] P. N. Srinivasu, N. Sandhya, R. H. Jhaveri, and R. Raut, "From Blackbox to Explainable AI in Healthcare: Existing Tools and Case Studies," *Mobile Information Systems*, vol. 2022, pp. 1–20, 2022, doi: 10.1155/2022/8167821.
[28] V. Vimbi, N. Shaffi, and M. Mahmud, "Interpreting artificial intelligence models: a systematic review on the application of LIME and SHAP in Alzheimer's disease detection," *Brain Informatics*, vol. 11, no. 1, 2024, doi: 10.1186/s40708-024-00222-1.
[29] M. H. Tilala *et al.*, "Ethical Considerations in the Use of Artificial Intelligence and Machine Learning in Health Care: A Comprehensive Review," *Cureus*, 2024, doi: 10.7759/cureus.62443.
[30] H. M. S. Alotaibi, W. Balachandran, and Z. Hunaiti, "Ethical Integration of AI in Healthcare Project Management: Islamic and Cultural Perspectives," *AI*, vol. 6, no. 12, p. 307, 2025, doi: 10.3390/ai6120307.
[31] R. Mustapha and S. N. A. Malkan, "Maqasid Al-Shariah in the AI Era: Balancing Innovation and Islamic Ethical Principles," *International Journal of Islamic Theology and Civilisation*, vol. 3, no. 3, pp. 1–21, 2025.
[32] F. Ali, K. Bouzoubaa, F. Gelli, B. Hamzi, and S. Khan, "Islamic Ethics and AI: An Evaluation of Existing Approaches to AI using Trusteeship Ethics," *Philosophy & Technology*, vol. 38, no. 3, 2025, doi: 10.1007/s13347-025-00922-4.
[33] M. M. Rahman, "Ensuring Halal Compliance in AI-Driven Healthcare Solutions," 2023, *IGI Global*. doi: 10.4018/979-8-3693-1082-3.ch015.
[34] A. Raquib, B. Channa, T. Zubair, and J. Qadir, "Islamic virtue-based ethics for artificial intelligence," *Discover Artificial Intelligence*, vol. 2, no. 1, 2022, doi: 10.1007/s44163-022-00028-2.
[35] A. S. Yunos, M. N. Hamdan, and M. Z. Ramle, "Artificial Intelligence in Healthcare: A Review from Islamic Perspectives," *Global Journal Al-Thaqafah*, vol. 15, no. 1, pp. 166–187, 2025, doi: 10.7187/GJAT072025-11.
[36] A. Al Kuwaiti *et al.*, "A Review of the Role of Artificial Intelligence in Healthcare," *Journal of Personalized Medicine*, vol. 13, no. 6, p. 951, 2023, doi: 10.3390/jpm13060951.
[37] A. M. Farid, F. A. Pratama, and D. Sulistyowati, "Maqashid Sharia Analysis on Personal Data Protection from the Dangers of Artificial Intelligence Misuse," *Jurnal Penelitian Ilmu-Ilmu Sosial*, vol. 6, no. 2, pp. 179–192, 2025, doi: 10.23917/sosial.v6i2.12636.

BIOGRAPHIES OF AUTHORS



Saeed Mubarak Abdullah Al-Shahrani holds a Bachelor's degree in Computer science from University of Bisha – Kingdom of Saudi Arabia, graduating with Excellent with Honors and a GPA of 4.54 out of 5. He also earned a Postgraduate Diploma in Computer Science Teaching Methods, which has contributed to enhancing his academic and educational expertise in computer technologies and education. He worked as an activity's supervisor at the University of Bisha for two years and is currently a teaching assistant at the same university in Saudi Arabia. He is also pursuing a Master of Philosophy in Computer Science at Universiti Teknologi Malaysia (UTM). His research interests focus on artificial intelligence, data science, machine learning, and explainable artificial intelligence (XAI), in addition to AI applications in healthcare, with a particular emphasis on ethical considerations and compliance with Islamic medical ethics. He can be contacted at email: Mubaraka@graduate.utm.my.



Haslina Binti Hashim is a senior lecturer at the Faculty of Computing, Universiti Teknologi Malaysia (UTM), Johor Bahru, Malaysia. She is a member of the Department of Applied Computing and Artificial Intelligence and is actively involved in the Intelligent Informatics Research Group (IIRG) and the UTM Big Data Centre (BDC). Her research interests include artificial intelligence, intelligent informatics, machine learning, data analytics, decision support systems, and big data applications. She has supervised numerous postgraduate students, participated in various national and international research grants, and has published scientific articles in reputable indexed journals and conference proceedings. Her current research focuses on developing intelligent AI-driven solutions for real-world applications, including healthcare and decision support systems. She can be contacted at email: haslinah@utm.my.