

Artificial Intelligence, Large Language Models, and Machine Learning in African Health Professions Education: A Systematic Review of Adoption, Literacy, Institutional Feasibility, and Ethical Perceptions

Promise Oladejo

oladejopromise@student.oauife.edu.ng

Obafemi Awolowo University

Oserebameh Augustine Beckley

Lagos State University College of Medicine, Lagos State University

John Chukwuemeka Amamdikwa

Stanford University

Joshua Oladiipo Owoyemi

University of Ilorin

Hans Johnson

University of Oxford

Systematic Review

Keywords: Artificial Intelligence, Large Language Models, Health Professions Education, Africa, Technology Acceptance, Medical Education

Posted Date: September 17th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10707419/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Background

Artificial intelligence (AI), including large language models (LLMs) and machine learning (ML), is increasingly present in health professions education (HPE), yet evidence on its adoption, literacy, institutional feasibility, and ethical implications specific to African settings remains fragmented. This review synthesizes evidence on AI/LLM/ML adoption, literacy, and institutional readiness among African health professions students, trainees, and educators.

Methods

This review followed PRISMA 2020 guidelines and was prospectively registered with PROSPERO (CRD420261356111). Searches were conducted in PubMed, Embase, and Scopus. Eligible studies examined AI/ML/LLM literacy, technology acceptance, adoption willingness, institutional feasibility, or curriculum integration among health professions students, trainees, or educators affiliated with African institutions. Study quality was appraised using an adapted Newcastle–Ottawa Scale, the Mixed Methods Appraisal Tool, the CASP qualitative checklist, and the JBI quasi-experimental checklist; certainty of evidence was evaluated using GRADE and GRADE-CERQual. A narrative synthesis was undertaken around five a priori thematic chapters.

Results

Of the 19 included studies, 16 were single-country African investigations, 2 were multinational surveys incorporating African student cohorts, and 1 was a non-African low- and middle-income-country comparator study evaluating institutional readiness. AI/LLM awareness was consistently high (54.8%–96%) but coexisted with minimal formal training (as low as 0.7% in one multinational cohort) and, in two studies using independently scored objective assessments, sharply overestimated self-rated competence. ChatGPT was the dominant tool named across the pool. Infrastructural constraints (unreliable electricity, high data costs, limited connectivity) were consistently identified as barriers, particularly in South African and Nigerian qualitative accounts. Ethical concerns and the institutional training vacuum were the two most richly evidenced themes (GRADE-CERQual: High confidence for both). Certainty of quantitative evidence for prevalence, predictors of use, and intervention effects ranged from Low to Very Low, reflecting single-study estimates, heterogeneous instruments, and, in one case, an absent control group.

Conclusion

AI/LLM adoption among African health professions students is widespread informally but is proceeding largely without institutional guidance, formal training, or infrastructure investment. Closing this gap through structured, ethically grounded curricular integration should be a priority for African HPE policy.

This review was registered with PROSPERO under registration number CRD420261356111.

Background

Artificial intelligence (AI), encompassing machine learning (ML) and, more recently, large language models (LLMs) such as ChatGPT, has rapidly transitioned from an emerging technology to an increasingly prominent component of health professions education (HPE). Following the public release of ChatGPT in November 2022 [1], student uptake accelerated worldwide, with widespread adoption occurring within months of its launch [2]. Health professions students are at the forefront of this transformation. In addition to mastering an expanding body of clinical knowledge, they are preparing for healthcare systems in which AI-enabled diagnostic support, administrative automation, research, and clinical decision-support tools are expected to become integral to professional practice.

Current evidence on AI adoption in HPE is derived predominantly from high-income countries, where reliable digital infrastructure, institutional AI policies, and formal digital literacy curricula facilitate implementation. African health professions institutions operate within markedly different contexts. Persistent challenges; including unreliable electricity supply, limited internet connectivity, high data costs, and minimal integration of AI into formal curricula, continue to constrain institutional adoption. Consequently, much of the AI use observed across the continent appears to be driven by individual initiative rather than structured educational implementation. Despite a rapidly expanding body of survey and qualitative research, it remains unclear whether this informal uptake is accompanied by adequate AI literacy, accurate self-assessment of competence, awareness of ethical considerations, or sufficient institutional support.

Although previous reviews have examined AI in medical education, most have adopted either a global perspective, in which African evidence is diluted within predominantly high-income settings, or a narrow focus on individual countries or professional groups. As a result, the regional evidence base remains fragmented, limiting understanding of the factors influencing AI adoption across African health professions education. To our knowledge, no previous systematic review has comprehensively synthesized evidence spanning medicine, nursing, pharmacy, and allied health professions across the African continent while integrating quantitative evidence on technology acceptance with qualitative accounts of student experiences and institutional contexts.

This review addresses that gap by systematically synthesizing evidence on AI, ML, and LLM adoption, literacy, institutional readiness, and ethical perceptions among African health professions students, trainees, and educators. The synthesis is guided by a three-layer conceptual framework that integrates the core constructs of the Unified Theory of Acceptance and Use of Technology (UTAUT): performance expectancy, effort expectancy, social influence, and facilitating conditions, with contextual moderators relevant to African health systems, including infrastructure and regulatory environments. The framework further distinguishes behavioural adoption from attitudinal and ethical responses, providing the conceptual foundation for interpreting the evidence presented in this review (Fig. 1).

Methods

Study design and registration

This systematic review was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 guidelines [3]. The review protocol was prospectively registered with the International Prospective Register of Systematic Reviews (PROSPERO; registration number CRD420261356111) [4].

Eligibility criteria (PICO-S)

Population (P): Undergraduate and postgraduate students enrolled in medicine, nursing, pharmacy, or allied health programmes, as well as early-career trainees (including house officers), affiliated with universities, tertiary institutions, or teaching hospitals in Africa.

Intervention/Domain of interest (I): Studies evaluating artificial intelligence (AI), machine learning (ML), or generative AI/large language models (LLMs) in relation to baseline literacy, perceptions, technology acceptance, willingness to adopt, educational needs, or curriculum integration. Studies evaluating conventional multimedia resources, standard e-learning platforms, or three-dimensional virtual reality (VR) simulators without AI-driven or machine learning functionality were excluded.

Comparator (C): Traditional teaching approaches, non-AI digital educational tools, existing educational pathways without AI integration, or comparator groups, where applicable.

Outcomes (O): Studies were required to report at least one of the following outcomes: (i) institutional readiness or feasibility, including infrastructure barriers, electricity interruptions, internet connectivity, data affordability, or the absence of institutional AI policies; or (ii) student-related outcomes, including AI literacy, knowledge, perceptions, technology acceptance, usage patterns, educational applications, ethical concerns, academic integrity issues, or predictors of AI adoption.

Study design (S): Cross-sectional, qualitative, mixed-methods, or quasi-experimental studies published as peer-reviewed articles in English.

Studies were eligible if they included undergraduate or postgraduate health professions trainees enrolled in African tertiary institutions. Multinational studies reporting African participant subgroups alongside non-African cohorts [5, 6] were retained to preserve relevant African data. One qualitative study conducted in a non-African low- and middle-income country (LMIC) [7] was also included as a contextual comparator because it explored institutional readiness and leadership within a resource-constrained academic health setting. In contrast, two non-African studies conducted in India (Selvi et al. [e9]) and Iran (Mousavi Baigi et al. [e10]) were excluded because they examined student knowledge and perceptions comparable to those synthesized from the African evidence base and did not provide additional contextual insights relevant to the review objectives.

Information sources

A literature search was conducted across three databases: PubMed, Embase, and Scopus. Databases were searched for articles published from 1 January 2015 up to 1 April 2026.

Search strategy

The search strategy combined three concept blocks: (1) AI/LLM/ML technology terms, (2) health professions education and population terms, and (3) African/low-resource setting terms, combined with the Boolean operator AND. Controlled vocabulary (MeSH/Emtree) terms were combined with free-text title/abstract/keyword terms specific to each database's syntax. The full search strategies for PubMed, Scopus, and Embase are provided in Additional file 1.

Study selection, data extraction, and conflict resolution

Records identified across the three databases (Embase $n = 371$; Scopus $n = 449$; PubMed $n = 121$; total $n = 941$) were screened for duplicates ($n = 262$ removed), leaving 679 records for title/abstract screening. Of these, 649 were excluded, leaving 30 reports sought for full-text retrieval; one report could not be retrieved, leaving 29 assessed for eligibility. Ten reports were excluded at the full-text stage (ineligible intervention $n = 4$; ineligible study design $n = 4$; ineligible population $n = 1$; ineligible study focus $n = 1$; full citations in Supplementary Table S1), leaving 19 studies included in the review. The complete study selection process is presented in Fig. 2, produced using the PRISMA2020 R package and Shiny application [8].

Title/abstract screening, full-text eligibility assessment, and data extraction were conducted independently by two reviewers (P.O. and O.B.). Data were extracted on study characteristics (author, year, country, study design), target population and setting, sample size, AI technology focus, infrastructure-related adaptations, and primary quantitative and qualitative outcomes. Any inter-rater discrepancies regarding study inclusion, extracted metrics, or quality appraisal classifications were initially addressed through iterative, consensus-driven discussion between the primary reviewers. Where consensus could not be reached, a third senior reviewer (J.O.) independently appraised the study to adjudicate the final decision.

Risk of bias assessment

Risk of bias was assessed according to study design. Cross-sectional studies ($n = 12$) were appraised using an adapted Newcastle–Ottawa Scale (NOS) for cross-sectional studies, evaluating Selection (maximum 5 stars), Comparability (maximum 1 star), and Outcome (maximum 3 stars) domains, for a maximum score of 9 [9]. Studies scoring 8–9 were classified as low risk of bias, 5–7 as moderate risk of bias, and 0–4 as high risk of bias. Qualitative studies ($n = 2$) were assessed using the Critical Appraisal Skills Programme (CASP) qualitative checklist [10]. Mixed-methods studies ($n = 3$) were appraised using the Mixed Methods Appraisal Tool (MMAT), which separately evaluates the qualitative and quantitative strands of a study before an overall integration judgement [11]. Quasi-experimental studies ($n = 2$) were

appraised using the Joanna Briggs Institute (JBI) critical appraisal checklist for quasi-experimental studies [12].

Certainty of evidence assessment

The certainty of evidence for quantitative outcomes was assessed using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach across five domains: risk of bias, inconsistency, indirectness, imprecision, and publication bias [13]. As all quantitative evidence derived from observational or non-randomized designs, certainty was initially rated as low and adjusted according to standard GRADE guidance. Confidence in qualitative synthesis findings was assessed using GRADE-CERQual across four domains: methodological limitations, coherence, adequacy of data, and relevance [14].

Data synthesis

Given substantial heterogeneity in study designs, populations, instruments, and outcome measures, a narrative synthesis was undertaken rather than meta-analysis. Findings were organized thematically around five chapters established a priori: (1) The Knowledge–Perception Paradox; (2) Informal Adoption and Tool Preference; (3) Experiential Deficit and Institutional Training Vacuum; (4) African Digital Divide and Infrastructural Constraints; and (5) Ethical Anxieties, Academic Integrity, and Professional Fears. Synthesis was further informed by a three-layer conceptual framework integrating UTAUT core constructs [15], African-specific contextual moderators, and a bifurcated behavioral/attitudinal outcome structure (Fig. 1).

Use of artificial intelligence tools

An AI-based large language model (Claude, Anthropic) was used during preparation of this manuscript to assist with drafting and formatting text, verifying reference details against source publications, and organising extracted data into tables and figures. The LLM was not used to conduct database searches, screen studies, extract primary data, or perform risk-of-bias, GRADE, or GRADE-CERQual appraisals; these steps were performed independently by the review team as described above. All AI-assisted content was reviewed and verified by the authors, who take full responsibility for the accuracy and integrity of the final manuscript.

Results

Search results

The electronic search identified 941 records across the three databases (Embase $n = 371$; Scopus $n = 449$; PubMed $n = 121$). After removal of 262 duplicates, 679 records were screened by title and abstract, of which 649 were excluded. Thirty reports were sought for full-text retrieval; one could not be retrieved, leaving 29 assessed for eligibility. Ten were excluded at full-text stage (ineligible intervention $n = 4$;

ineligible study design n = 4; ineligible population n = 1; ineligible study focus n = 1), leaving 19 studies included in the review (Fig. 2).

Geographic distribution of included studies

Figure 3 summarizes the geographic distribution of the 16 single-country included studies by approximate country centroid, with marker size and color indicating the number of contributing studies per country. Nigeria and Egypt each contributed five studies, the largest single-country totals in the pool, followed by Ghana (two studies) and single studies each from Morocco, South Africa, Sudan, and Uganda. Three further cohorts extend beyond a single African country [5–7] and are described separately in Fig. 3's note rather than plotted, since their populations are not confined to a single country centroid.

Characteristics of included studies

Of the 19 included studies, 16 were single-country African investigations (Ghana n = 2, Nigeria n = 5, Egypt n = 5, Morocco n = 1, South Africa n = 1, Sudan n = 1, Uganda n = 1), 2 were multinational surveys incorporating African student cohorts (Ejaz et al. [5], a 48-country IFMSA cohort; Elnaem et al. [6], a 9-country pharmacy-education cohort), and 1 was a non-African LMIC comparator study evaluating institutional readiness and leadership dynamics (Jafri et al. [7], Pakistan), retained per the eligibility criteria set out in Methods. Study designs comprised cross-sectional surveys (n = 12), mixed-methods studies (n = 3), quasi-experimental studies (n = 2), and qualitative studies (n = 2). Sample sizes ranged from small qualitative interview/focus-group cohorts up to 2,009 respondents in the largest cross-national survey. ChatGPT was the dominant or sole AI technology focus in the majority of included studies. Full study characteristics are presented in Table 1.

Table 1

Characteristics of included studies (n = 19). Asterisks denote the three cohorts extending beyond a single African country (see Methods, Eligibility criteria).

#	Author, Year	Country	Design	Population	N	AI Focus
1	Ankamah et al., 2024	Ghana	Cross-sectional	Medical students	123	General AI tools
2	David-Olawade et al., 2025	Nigeria	Cross-sectional	Students & staff, OAU	551	General AI literacy
3	Olawade et al., 2025	Nigeria	Cross-sectional	Nursing students, 5 univ.	676	General AI literacy
4	Chakri et al., 2025	Morocco	Cross-sectional	Medical students	580	General AI
5	Ghanem et al., 2025	Egypt	Cross-sectional	Medical students, 10 univ.	423	Generative AI
6	Allam et al., 2024	Egypt	Cross-sectional	Medical students & house officers	1346	AI (general)
7	Abdelhafiz et al., 2025	Egypt	Cross-sectional	Medical students	614	AI (general)
8	Oluwadiya et al., 2023	Nigeria	Cross-sectional	Students & lecturers, 10 schools	708 (+ 294 lecturers)	AI (general)
9	Ajalo et al., 2025	Uganda	Cross-sectional	Medical students, 4 univ.	564	ChatGPT / AI tools
10	Elnaem et al., 2025	9 countries*	Cross-sectional	Pharmacy students	2009	Generative AI (UTAUT)
11	Ahmed et al., 2025	Sudan	Cross-sectional	Medical students	1443	ChatGPT
12	Orok et al., 2024	Nigeria	Cross-sectional	Pharmacy students, ABUAD	252	Chat-based AI
13	Iliyasu et al., 2024	Nigeria	Mixed-methods	Medical students, Bayero Univ.	420	ChatGPT
14	Eltaybani et al., 2025	Egypt	Mixed-methods	Nursing staff & students	—	LLMs (general)
15	Ejaz et al., 2023	48 countries*	Mixed-methods	Medical students (IFMSA)	—	AI (general)

#	Author, Year	Country	Design	Population	N	AI Focus
16	Abdelmohsen et al., 2026	Egypt	Quasi-experimental	Undergrad. & junior residents, Al-Azhar	240	AI-enhanced video content
17	Donkor et al., 2025	Ghana	Quasi-experimental	Medical imaging students, KNUST	144	AI in medical imaging
18	Jafri et al., 2026	Pakistan*	Qualitative	Healthcare leadership	—	AI (institutional)
19	Lewis et al., 2024	South Africa	Qualitative	Radiation sciences students	—	AI tools (general)

Risk of bias / methodological quality

Overall, 6 of 19 studies (32%) were judged at low risk of bias, 11 (58%) at moderate risk, and 2 (11%) at high risk (Fig. 4). Among the 12 cross-sectional studies appraised with the adapted Newcastle–Ottawa Scale, only one (David-Olawade et al. [16]) achieved a low-risk rating (8/9), reflecting its large sample, dual-reviewer objective knowledge scoring, and robust statistical adjustment; nine were rated moderate risk, and two (Ankamah et al. [17] and Orok et al. [18]) were rated high risk, principally reflecting the absence of validated instruments, absent or minimal multivariable adjustment, and unclear or unreported non-response handling. All three mixed-methods studies (Iliyasu et al. [19]; Eltaybani et al. [20]; Ejaz et al. [5]) were rated high quality (low risk of bias) on the Mixed Methods Appraisal Tool. Both qualitative studies (Jafri et al. [7]; Lewis et al. [21]) were rated low risk of bias on the CASP checklist, with Jafri et al. [7] meeting all ten CASP criteria. Both quasi-experimental studies (Abdelmohsen et al. [22]; Donkor et al. [23]) were rated moderate risk of bias on the JBI checklist, in both cases principally due to the absence of a control group or a post-test-only design. Full risk-of-bias appraisals are presented in Table 2 and summarized graphically in Fig. 4.

Table 2
Risk of bias / methodological quality summary (n = 19).

#	Author, Year	Appraisal Tool	Score / Verdict	Risk of Bias
1	Ankamah et al., 2024	NOS	3/9	High
2	David-Olawade et al., 2025	NOS	8/9	Low
3	Olawade et al., 2025	NOS	7/9	Moderate
4	Chakri et al., 2025	NOS	5/9	Moderate
5	Ghanem et al., 2025	NOS	6/9	Moderate
6	Allam et al., 2024	NOS	5/9	Moderate
7	Abdelhafiz et al., 2025	NOS	6/9	Moderate
8	Oluwadiya et al., 2023	NOS	6/9	Moderate
9	Ajalo et al., 2025	NOS	7/9	Moderate
10	Elnaem et al., 2025	NOS	6/9	Moderate
11	Ahmed et al., 2025	NOS	5/9	Moderate
12	Orok et al., 2024	NOS	4/9	High
13	Iliyasu et al., 2024	MMAT	High quality (both strands)	Low
14	Eltaybani et al., 2025	MMAT	High quality (qual-dominant)	Low
15	Ejaz et al., 2023	MMAT	High quality (both strands)	Low
16	Abdelmohsen et al., 2026	JBIR	7 Yes / 1 No / 1 Partial	Moderate
17	Donkor et al., 2025	JBIR	7 Yes / 1 No / 1 Partial	Moderate
18	Jafri et al., 2026	CASP	10/10	Low
19	Lewis et al., 2024	CASP	9/10	Low

Certainty of evidence

Certainty of quantitative evidence, assessed using GRADE, ranged from Low to Very Low across five outcome domains (Table 3). No outcome achieved Moderate or High certainty, reflecting the pool's reliance on observational designs, frequent single-study estimates for specific effect sizes, and serious inconsistency in how core constructs (e.g., "awareness," "use") were operationalized across studies. Confidence in qualitative synthesis findings, assessed using GRADE-CERQual, was higher: two of the five thematic chapters (Experiential Deficit & Institutional Training Vacuum; Ethical Anxieties, Academic Integrity, and Professional Fears) achieved High confidence, while the remaining three achieved

Moderate confidence, most often limited by thinner supporting data on a specific sub-claim within an otherwise well-evidenced theme rather than by any concern affecting the theme as a whole (Table 4).

Table 3
GRADE certainty of quantitative evidence, by outcome.

Outcome	Certainty	Primary downgrade reasons
1. AI/LLM awareness prevalence	Low	Inconsistency, indirectness
2. AI/LLM use/adoption prevalence	Low	Inconsistency, indirectness
3. Predictors of adoption (aORs)	Low	Indirectness, imprecision (single study)
4. AI-enhanced video → clarity/knowledge	Low	Risk of bias, indirectness (offset partly by large effect)
5. AI-ethics curriculum → knowledge score	Very Low	Risk of bias (no control group), indirectness

Table 4
GRADE-CERQual confidence in qualitative synthesis findings, by theme.

Theme	Confidence	Primary limiting factor
1. Knowledge–Perception Paradox	Moderate	Coherence (miscalibration sub-claim: Nigeria-only)
2. Informal Adoption & Tool Preference	Moderate	Adequacy of data (motivators sub-claim)
3. Experiential Deficit & Training Vacuum	High	—
4. African Digital Divide & Infrastructure	Moderate	Adequacy of data (single rich source)
5. Ethical Anxieties & Professional Fears	High	—

Thematic synthesis

Chapter 1: The knowledge–perception paradox

Across the included studies, awareness of AI and LLMs was consistently high among African health professions students, with reported awareness ranging from 54.8% [18] to 96% [24]. However, this widespread familiarity did not translate into equivalent levels of competence or formal educational preparation. Instead, the evidence consistently revealed a knowledge–perception paradox, whereby students reported confidence in their understanding of AI despite limited objective knowledge and minimal institutional training.

The strongest evidence for this discrepancy came from two Nigerian studies that independently assessed AI knowledge using the same objective scoring instrument. Although a substantial proportion of participants rated their AI knowledge positively, objective assessments demonstrated that only a very small minority achieved high knowledge scores, while the overwhelming majority were classified as having low objective knowledge [16, 25]. Self-perceived competence was significantly associated with objective performance, yet consistently overestimated actual knowledge, indicating systematic miscalibration rather than accurate self-appraisal.

A related pattern emerged across studies examining knowledge, AI use, and educational exposure. Students generally demonstrated moderate awareness of AI technologies and reported widespread familiarity with tools such as ChatGPT, yet regular use remained less common and formal instruction was exceedingly rare [19]. In the multinational IFMSA survey, fewer than 1% of respondents reported receiving substantial formal AI education despite most describing themselves as having at least moderate AI knowledge [5]. Similarly, students who were aware of LLMs frequently reported little or no practical experience, with uncertainty regarding effective use representing a persistent barrier even among those familiar with the technology [20].

Collectively, these findings suggest that AI knowledge among African health professions students has developed primarily through self-directed learning rather than structured educational programmes. They also indicate that self-reported knowledge should be interpreted cautiously, as perceived competence frequently exceeds objectively measured understanding. Although evidence for inaccurate self-assessment is currently derived from Nigerian studies, the broader disconnect between awareness, practical use, and formal training was consistently observed across multiple countries and educational settings (GRADE-CERQual: Moderate confidence).

Chapter 2: Informal adoption and tool preference

Across the included studies, ChatGPT emerged as the dominant AI application used by African health professions students. Wherever specific AI tools were compared, ChatGPT consistently ranked first in both recognition and reported use, substantially exceeding alternative platforms such as Microsoft Copilot, Google Gemini, Claude, Bing AI, and Snapchat AI [2, 20]. This pattern was observed across diverse educational settings, including Nigeria, Egypt, Sudan, Uganda, and multinational cohorts, suggesting that early AI adoption within African health professions education has been driven largely by a single, highly accessible platform.

Evidence explaining why students favoured ChatGPT was comparatively limited. The only study to examine motivations directly identified ease of use and time efficiency as the principal drivers of adoption, followed by peer recommendation, perceived accuracy, and trust in the technology [20]. These findings suggest that practical utility, rather than institutional endorsement or curricular integration, has largely shaped students' decisions to engage with AI.

Patterns of AI use were remarkably consistent across studies. Academic activities: including assignment completion, examination preparation, tutorial support, literature searching, and research writing, represented the predominant applications of AI tools [2]. Non-academic uses, such as emotional support, companionship, and recreational engagement, were also reported but occurred less frequently. Overall, the evidence indicates that AI has been incorporated primarily as an informal academic support tool, with adoption occurring largely outside structured educational frameworks (GRADE-CERQual: Moderate confidence).

Chapter 3: Experiential deficit and the institutional training vacuum

The absence of formal AI education emerged as one of the most consistent findings across the included studies (GRADE-CERQual: High confidence). Across diverse educational settings, students reported widespread exposure to AI tools despite receiving little or no structured instruction in their use. In Ghana, all participants in an educational intervention study reported having received no prior training in medical imaging AI before the intervention was delivered [23]. Similarly, the multinational IFMSA survey found that fewer than 1% of respondents had received substantial formal AI education, despite overwhelming support for incorporating AI into undergraduate curricula and for interdisciplinary teaching involving computer science and biomedical engineering [5].

The training deficit extended beyond students to institutional capacity. Qualitative evidence indicated that limited AI expertise among faculty, financial constraints, and difficulties procuring AI technologies restricted institutions' ability to implement AI education effectively [7]. Together, these findings suggest that the current educational gap reflects broader institutional limitations rather than deficiencies in student motivation or interest. Across the available evidence, AI adoption has consistently outpaced institutional preparedness.

Chapter 4: African digital divide and infrastructural constraints

Infrastructure-related barriers were reported across multiple studies, although the depth of evidence varied between settings (GRADE-CERQual: Moderate confidence). Students consistently described unreliable electricity supply, unstable internet connectivity, high data costs, and limited access to high-performance computing resources as important obstacles to engaging with AI-enabled educational technologies [21]. These barriers affected both access to AI applications and the reliability with which students could integrate them into routine learning.

Evidence from quantitative studies supported these qualitative findings. Students frequently perceived institutional digital infrastructure as inadequate for AI-enabled education, while several studies identified internet access and limited institutional resources as persistent constraints [16, 17]. Although some authors highlighted investments in electricity and broadband connectivity as priorities for national policy, these recommendations were generally based on interpretation rather than directly measured outcomes [23].

Notably, one study presented a contrasting perspective. Rather than representing a barrier to AI implementation, inadequate physical infrastructure prompted the use of AI as an alternative solution. In an Egyptian quasi-experimental study, AI-assisted background generation successfully replaced the need for a dedicated green-screen studio during educational video production, substantially reducing production costs while maintaining instructional quality [22]. This finding suggests that, in some settings, AI may function not only as a technology constrained by limited infrastructure but also as a practical means of overcoming infrastructural limitations.

Chapter 5: Ethical anxieties, academic integrity, and professional concerns

Ethical concerns surrounding AI emerged as one of the most consistently reported themes across the included studies (GRADE-CERQual: High confidence). Students expressed concerns spanning academic integrity, declining clinical reasoning, misinformation, professional identity, data privacy, and the broader implications of AI for healthcare practice. Importantly, these concerns persisted despite generally favourable attitudes towards AI adoption, indicating that acceptance and apprehension frequently coexisted rather than representing opposing viewpoints.

One particularly noteworthy finding was that greater AI knowledge did not necessarily reduce ethical concerns. In a multicentre Nigerian study comparing students with lecturers, students demonstrated significantly higher AI knowledge while simultaneously expressing greater concern that AI could dehumanize patient care, reduce physicians' clinical skills, and ultimately threaten professional roles [26]. Rather than resolving uncertainty, greater familiarity with AI appeared to coincide with increased awareness of its potential risks.

Direct evidence of academic integrity challenges was also identified. Participants in one qualitative study openly acknowledged using LLMs to complete online examinations and verify answers before submission [20]. The authors attributed these unusually candid disclosures to methodological choices that fostered participant trust, highlighting the value of qualitative approaches for exploring behaviours that are unlikely to be captured through conventional surveys.

Quantitative findings reinforced these concerns. Students frequently identified overreliance on AI, deterioration of critical thinking and clinical reasoning, inaccurate or fabricated information, unreliable outputs, fabricated references, and academic dishonesty as important risks associated with AI use [18, 19]. Concerns regarding responsible AI use, self-regulation, and academic integrity were consistently reported across studies from Nigeria, Egypt, and Sudan [27–30].

Beyond educational settings, concerns extended to institutional and professional practice. Participants described uncertainty regarding future professional roles, the possibility of excessive reliance on AI within healthcare systems, and broader questions surrounding workforce preparedness [7]. Students also demonstrated an awareness of emerging detection practices, with some reporting deliberate modification of AI-generated work to avoid appearing implausibly sophisticated [21]. Collectively, these findings suggest that African health professions students view AI not only as an educational resource

but also as a technology that raises complex ethical and professional questions requiring explicit curricular attention.

Discussion

Statement of key findings

This review synthesized evidence from 19 studies examining AI, ML, and LLM adoption, literacy, institutional readiness, and ethical perceptions among African health professions students, trainees, and educators. Four principal findings emerged. First, although awareness of AI was consistently high, it frequently exceeded objectively measured competence and was rarely supported by formal educational preparation, giving rise to a recurring knowledge–perception paradox. Second, AI adoption was predominantly informal and centred on ChatGPT, with engagement driven largely by ease of use, accessibility, and peer influence rather than institutional guidance. Third, the absence of structured AI education was consistently reported across countries, professions, and study designs, indicating a widespread institutional training gap. Finally, ethical concerns—including academic integrity, overreliance on AI, and uncertainty regarding the future of professional practice—were among the most consistently reported themes across the evidence base.

Interpretation in context

The knowledge–perception paradox identified in this review aligns with broader evidence from technology adoption research, in which self-perceived digital competence often exceeds objectively demonstrated ability, particularly for rapidly evolving technologies lacking established educational pathways. The present review extends this literature by demonstrating that inaccurate self-assessment occurs within an educational environment where opportunities for formal training are largely absent. Without structured instruction or objective assessment, students are required to develop AI competencies independently, limiting opportunities to calibrate perceived competence against measurable standards.

The consistency of the institutional training gap across geographically and methodologically diverse studies suggests that it reflects a broader regional challenge rather than isolated institutional shortcomings. Similar findings were observed across studies from Ghana, Egypt, Nigeria, and multinational cohorts, regardless of whether evidence was generated through cross-sectional surveys, qualitative inquiry, or mixed-methods designs. Collectively, these findings indicate that student adoption of AI is advancing more rapidly than institutional efforts to integrate AI into curricula, faculty development, and educational policy.

The review also highlights a more nuanced relationship between infrastructure and AI adoption than is often assumed. Although unreliable electricity, limited internet connectivity, and high data costs were consistently identified as barriers, one study demonstrated that AI itself could compensate for infrastructure limitations by replacing resource-intensive educational processes with lower-cost

alternatives. This suggests that infrastructure and AI adoption should not necessarily be viewed as sequential prerequisites. Instead, AI may, in some contexts, represent part of the solution to existing resource constraints.

Implications for policy and practice

The findings suggest that the principal challenge for African health professions education is no longer encouraging students to adopt AI, but ensuring that this widespread adoption is accompanied by appropriate educational support, institutional governance, and ethical guidance. Across the included studies, students were already incorporating AI into their academic activities despite receiving little formal instruction in its appropriate use.

These findings support the integration of structured AI literacy and AI ethics within undergraduate health professions curricula. Such initiatives should extend beyond technical proficiency to include critical appraisal of AI-generated information, responsible academic use, data governance, patient confidentiality, bias, and the limitations of AI in clinical decision-making. Importantly, curriculum development should be proactive rather than reactive, recognising that student adoption has already outpaced institutional preparedness.

Strengths and limitations

A major strength of this review is the breadth and methodological diversity of the evidence synthesized. The review included cross-sectional, qualitative, mixed-methods, and quasi-experimental studies spanning multiple African countries, with methodological quality assessed using design-specific appraisal tools and certainty of evidence evaluated using both GRADE and GRADE-CERQual. Integrating quantitative and qualitative evidence enabled triangulation across complementary sources, strengthening confidence in findings that were consistently observed across different methodological approaches.

Several limitations should also be considered. Most quantitative findings were supported by low- or very low-certainty evidence because of the observational nature of the included studies, methodological heterogeneity, and inconsistency in the measurement of key constructs such as AI awareness, literacy, and use. These differences limited direct quantitative comparison across studies and precluded meta-analysis. The geographical distribution of evidence was also uneven, with Nigeria and Egypt accounting for a substantial proportion of the included studies, while Francophone Africa, Central Africa, and much of Southern Africa remained underrepresented. Finally, although three included studies extended beyond single-country African populations, they were retained because they provided relevant African participant data or valuable contextual comparisons for resource-constrained educational settings.

Future research

Future research should prioritise the development and validation of objective measures of AI literacy to determine whether the knowledge–perception paradox identified in Nigerian studies is observed across

other African contexts. Longitudinal and controlled intervention studies are also needed to evaluate the educational impact of formal AI curricula and strengthen the currently limited causal evidence base. Greater qualitative exploration of students' motivations, patterns of AI use, and decision-making processes would improve understanding of why particular AI tools are adopted and how they influence learning behaviours. Finally, expanding research across currently underrepresented regions of Africa will be essential to improve the geographical representativeness of future evidence syntheses.

Conclusion

AI adoption among African health professions students is widespread and is occurring primarily through informal, student-driven engagement rather than structured institutional implementation. Although awareness of AI is high, formal educational preparation remains limited, creating a persistent gap between rapid technological adoption and institutional readiness. At the same time, students demonstrate substantial awareness of the ethical and professional challenges associated with AI, underscoring the need for educational approaches that address responsible use alongside technical competence.

Collectively, these findings suggest that integrating structured AI literacy and AI ethics into health professions education should be considered an important priority for curriculum development across African institutions. Supporting students with evidence-informed training, appropriate governance, and institutional guidance will be essential to ensure that increasing AI adoption translates into safe, ethical, and effective educational and clinical practice.

Abbreviations

AI	
Artificial Intelligence	
LLM	
Large Language Model	
ML	
Machine Learning	
HPE	
Health Professions Education	
PRISMA	
Preferred Reporting Items for Systematic Reviews and Meta-Analyses	
PROSPERO	
International Prospective Register of Systematic Reviews	
NOS	
Newcastle–Ottawa Scale	
CASP	
Critical Appraisal Skills Programme	

MMAT

Mixed Methods Appraisal Tool

JB

Joanna Briggs Institute

GRADE

Grading of Recommendations Assessment, Development and Evaluation

CERQual

Confidence in the Evidence from Reviews of Qualitative research

UTAUT

Unified Theory of Acceptance and Use of Technology

TAM

Technology Acceptance Model

CI

Confidence Interval

OR

Odds Ratio

aOR

Adjusted Odds Ratio

IFMSA

International Federation of Medical Students' Associations

LMIC

Low- and Middle-Income Country.

Declarations

Ethics approval and consent to participate

This study was a systematic review of published literature and did not involve human participants or the collection of primary data. Ethical approval was therefore not required.

Consent for publication

Not applicable.

Availability of data and materials

All data generated or analysed during this study are derived from previously published sources. The complete list of references and extracted data tables supporting the findings are available within the manuscript. The study protocol is available on PROSPERO at <https://www.crd.york.ac.uk/PROSPERO/view/CRD420261356111>.

Competing interests

The authors declare that they have no competing interests.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Authors' contributions

PO conceived the review, developed the methodology, registered the protocol, led title/abstract screening, full-text eligibility assessment, data extraction, and methodological quality appraisal, performed the formal analysis, produced the visualisations, and drafted the original manuscript. OB independently conducted title/abstract screening, full-text eligibility assessment, data extraction, and methodological quality appraisal, contributed to formal analysis, and reviewed and edited the manuscript. JA contributed to conceptualization and methodological oversight, and critically reviewed and edited the manuscript. JO adjudicated discrepancies in screening, extraction, and risk-of-bias appraisal as third reviewer, validated the data, and reviewed and edited the manuscript. HJ contributed to conceptualization and methodological oversight, critically reviewed and edited the manuscript, provided overall supervision, and served as review guarantor.

All authors read and approved the final manuscript.

Acknowledgements

We acknowledge the authors of the studies used in this review.

References

1. Introducing ChatGPT [Internet]. OpenAI. 2022 Nov [cited 2026 Aug 13]. Available from: <https://openai.com/index/chatgpt/>
2. Ajalo E, Mukunya D, Nantale R, Kayemba F, Pangholi K, Babuya J, Langoya Akuu S, Namiiro AM, Nsubuga YB, Mpagi JL, Musaba MW. Widespread use of ChatGPT and other Artificial Intelligence tools among medical students in Uganda: A cross-sectional study. PLoS ONE. 2025;20(1):e0313776.
3. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, Shamseer L, Tetzlaff JM, Akl EA, Brennan SE, Chou R. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;372:n71.

4. Oladejo P, Beckley OA, Amamdikwa JC, Owoyemi JO, Johnson H, Artificial Intelligence. Large Language Models, and Machine Learning in African Health Professions Education: A Systematic Review of Adoption, Literacy, Institutional Feasibility, and Ethical Perceptions. PROSPERO 2026 CRD420261356111. Available from: <https://www.crd.york.ac.uk/PROSPERO/view/CRD420261356111>
5. Ejaz H, McGrath H, Wong BL, Guise A, Vercauteren T, Shapey J. Artificial intelligence and medical education: A global mixed-methods study of medical students' perspectives. *Digit Health*. 2022;8:20552076221089099.
6. Elnaem MH, Okuyan B, Mubarak N, Thabit AK, AbouKhatwa MM, Ramatillah DL, Isah A, Al-Jumaili AA, Nazar NI. Students' acceptance and use of generative AI in pharmacy education: international cross-sectional survey based on the extended unified theory of acceptance and use of technology. *Int J Clin Pharm*. 2025;47(4):1097–108.
7. Jafri L, Majid H, Khan AH, Kapadia A, Rehman S, Hilal K, Siddiqui I, Farooqui A, Ghias K, Rashid N, Sarfaraz Y. Evidence-based action plan for integrating artificial intelligence in an academic medical centre-a multidisciplinary approach. *PLoS ONE*. 2026;21(4):e0346582.
8. Haddaway NR, Page MJ, Pritchard CC, McGuinness LA. PRISMA2020: An R package and Shiny app for producing PRISMA 2020-compliant flow diagrams, with interactivity for optimised digital transparency and Open Synthesis. *Campbell Syst Rev*. 2022;18:e1230. 10.1002/cl2.1230.
9. Gualdi-Russo E, Zaccagni L. The Newcastle–Ottawa Scale for Assessing the Quality of Studies in Systematic Reviews. *Publications*. 2026;14(1):4. 10.3390/publications14010004.
10. Critical Appraisal Skills Programme (CASP). CASP Qualitative Checklist [Internet]. Oxford: CASP. 2018 [cited 2026 Aug 13]. Available from: <https://casp-uk.net/casp-tools-checklists/>
11. Hong QN, Fàbregues S, Bartlett G, Boardman F, Cargo M, Dagenais P, et al. The Mixed Methods Appraisal Tool (MMAT) version 2018 for information professionals and researchers. Montreal: McGill University; 2018.
12. Tufanaru C, Munn Z, Aromataris E, Campbell J, Hopp L. Chapter 3: Systematic reviews of effectiveness. In: Aromataris E, Munn Z, editors. *JBIManual for Evidence Synthesis* [Internet]. Adelaide: JBI; 2020. Available from: <https://synthesismanual.jbi.global/>
13. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ*. 2008;336:924–6. 10.1136/bmj.39489.470347.AD.
14. Lewin S, Glenton C, Munthe-Kaas H, Carlsen B, Colvin CJ, Gülmezoglu M, et al. Using qualitative evidence in decision making for health and social interventions: an approach to assess confidence in findings from qualitative evidence syntheses (GRADE-CERQual). *PLoS Med*. 2015;12(10):e1001895. 10.1371/journal.pmed.1001895.
15. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: Toward a unified view. *MIS Q*. 2003;27(3):425–78.

16. David-Olawade AC, Wada OZ, Adeniji YJ, Aderupoko IV, Olawade DB. Artificial intelligence readiness among healthcare students in Nigeria: A cross-sectional study assessing knowledge gaps, exposure, and adoption willingness. *Int J Med Inf.* 2025;204:106085. 10.1016/j.ijmedinf.2025.106085.
17. Ankamah S, Gyesi K, Amponsah V. Awareness, knowledge, and attitude towards artificial intelligence: Perspective of medical students in Ghana. *Inf Dev.* 2025;41(3):843–58.
18. Orok E, Okaramée C, Egboro B, Egbochukwu E, Bello K, Etukudo S, Ogologo MS, Onyeka P, Etukokwu O, Kolawole M, Orire A. Pharmacy students' perception and knowledge of chat-based artificial intelligence tools at a Nigerian University. *BMC Med Educ.* 2024;24(1):1237.
19. Iliyasu Z, Abdullahi HO, Iliyasu BZ, Bashir HA, Amole TG, Abdullahi HM, Abdullahi AU, Kwaku AA, Dahir T, Tsiga-Ahmed FI, Jibo AM. Correlates of medical and allied health students' engagement with generative AI in Nigeria. *Med Sci Educ.* 2024;35(1):269–80.
20. Eltaybani S, Ali HF, Abdelhalim GE. Exploring nurse educators' and students' use of large language models for academic purposes in a developing country: Unethical uses reported. *Nurse Educ Pract.* 2025 Aug;6:104502.
21. Lewis S, Bhyat F, Casmod Y, Gani A, Gumede L, Hajat A, Hazell L, Kammies C, Mahlaola TB, Mokoena L, Vermeulen L. Medical imaging and radiation science students' use of artificial intelligence for learning and assessment. *Radiography.* 2024;30:60–6.
22. Abdelmohsen A, Elshehawy N, Abdel Razek M, Badawy R, Sharaf S, Abdelmohsen O, Abdelmohsen H, Abdelmohsen M. Educational impact and cost efficiency of AI-enhanced videos in pediatric surgery training: a quasi-experimental study. *Sci Rep.* 2026;16(1):8952.
23. Donkor A, Boakye E, Atuanor P, Wiafe YA. Evaluation of a classroom-based medical imaging artificial intelligence educational intervention in Ghana: a pre-test/post-test study design. *Radiography.* 2025;31(4):102987.
24. Chakri I, El Khayali O, Lahlou L. Knowledge and perceptions of AI among medical students in Morocco: cross-sectional study. *JMIR Form Res.* 2025;9(1):e66156.
25. Olawade DB, David-Olawade AC, Rotifa OB, Wada OZ. Artificial intelligence in Nigerian nursing education: Are future nurses prepared for the digital revolution in healthcare? *Nurse Educ Pract.* 2025 Aug 13:104511.
26. Oluwadiya KS, Adeoti AO, Agodirin SO, Nottidge TE, Usman MI, Gali MB, Onyemaechi NO, Ramat AM, Adedire A, Zakari LY. Exploring artificial intelligence in the Nigerian medical educational space: An online cross-sectional study of perceptions, risks and benefits among students and lecturers from ten universities. *Niger Postgrad Med J.* 2023;30(4):285–92.
27. Ghanem OA, Hagag AM, Kormod ME, El-Refaay MA, Khedr AM, Abozaid OM, Abdelmoaty KM, Hamed MS. Medical students' knowledge, attitudes, and practices toward generative artificial intelligence in Egypt 2024: a Cross-Sectional study. *BMC Med Educ.* 2025;25(1):790.
28. Allam RM, Abdelfatah D, Khalil MI, Elsaieed MM, El Desouky ED. Medical students and house officers' perception, attitude and potential barriers towards artificial intelligence in Egypt, cross

29. Abdelhafiz AS, Farghly MI, Sultan EA, Abouelmagd ME, Ashmawy Y, Elsebaie EH. Medical students and ChatGPT: analyzing attitudes, practices, and academic perceptions. BMC Med Educ. 2025;25(1):187.
30. Ahmed WM, Abdalmotalib MM, Elbadawi MH, Mohammed GT, Mohamed WM, Mohammed FS, Salih HS, Mohamed HO. Shaping the future of medical education: A cross-sectional study on ChatGPT attitude and usage among medical students in Sudan. PLoS ONE. 2025;20(5):e0324698.

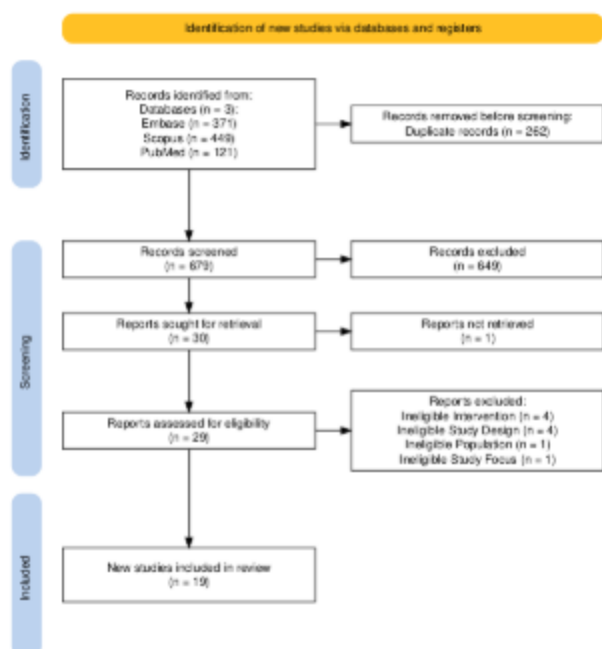


Figure 2

PRISMA 2020 flow diagram of study identification, screening, and inclusion, produced using the PRISMA2020 R package and Shiny application [8].

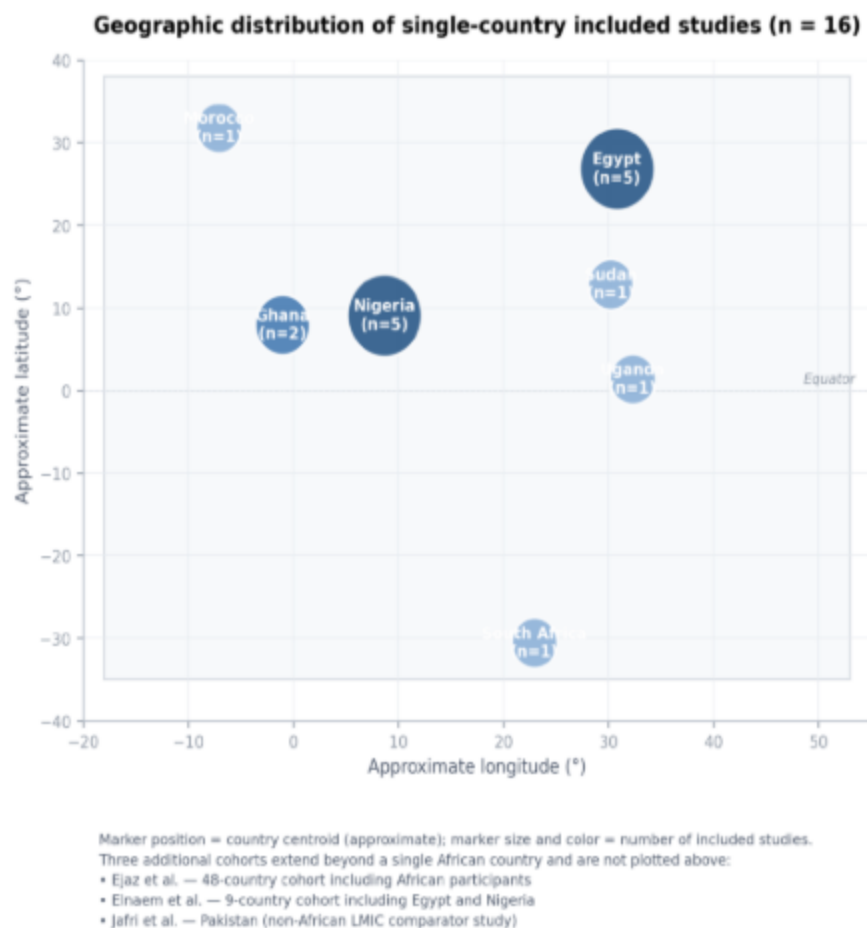


Figure 3

Geographic distribution of the 16 single-country included studies, by approximate country centroid.

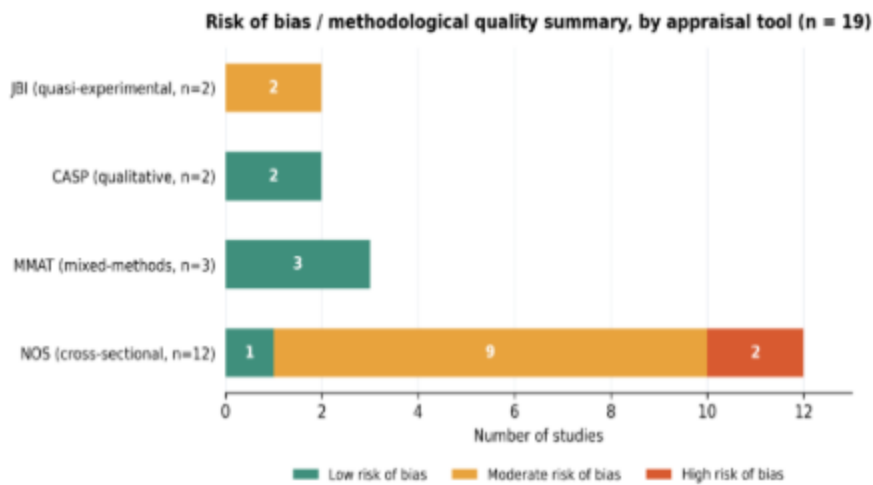


Figure 4

Risk of bias / methodological quality summary, by appraisal tool (n = 19).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Supplementaryinfo.docx](#)