Viewpoint

Augmented With AI: Practical Guide for Clinician Educators Engaged in Health Professions Education Scholarship

Zareen Zaidi[1*], MD, PhD; Rohini Ganjoo[2*], PhD; Kirsten Brown[3*], PhD; Farida Izzi[1*], MD; Anna Rubin[1*], MD; Ioannis Koutroulis[4,5*], MD, PhD

[1]Department of Medicine, School of Medicine and Health Sciences, George Washington University, Washington, DC, United States
[2]Department of Biomedical Laboratory Sciences, School of Medicine and Health Sciences, George Washington University, Washington, DC, United States
[3]Department of Anatomy and Cell Biology, School of Medicine and Health Sciences, George Washington University, Washington, DC, United States
[4]Trustworthy AI Initiative, George Washington University, Washington, DC, United States
[5]Sheikh Zayed Center for Surgical Innovation, Children's National Research Institute, Washington, DC, United States
*all authors contributed equally

Corresponding Author:

Zareen Zaidi, MD, PhD
Department of Medicine, School of Medicine and Health Sciences
George Washington University
2150 Pennsylvania Ave. NW Suite 5-416N
Washington, DC 200037
United States
Phone: 1 202-741-2222 ext 48218
Fax: 1 202-741-2222
Email: zzaidi@email.gwu.edu

Abstract

Clinician educators occupy a demanding role at the intersection of clinical practice, teaching, and scholarship. Given their teaching roles, they are expected to contribute in particular to health professions education scholarship, yet many have little protected time. AI tools have emerged as practical support for 2 of the time-intensive aspects of scholarly activities: conducting early-stage literature reviews and improving scholarly writing. The rapidly expanding AI landscape has made it increasingly difficult to distinguish purpose-built scholarly platforms from general-purpose AI assistants. As a result, many clinician educators remain uncertain about which tools are appropriate for academic work and how they can be used responsibly. Additionally, the barriers are often as much institutional as individual. Limited protected time, access to experienced mentors, methodological expertise, editorial support, and library services all shape the ability of clinician educators to engage in scholarship. Although AI tools may reduce some of the burdens associated with literature review and writing, they should not be viewed as substitutes for these institutional supports. This viewpoint is intended for clinician educators seeking practical guidance on integrating AI into health professions education scholarship. Our aim is not to provide a systematic review, an exhaustive list of platforms, or an endorsement or ranking of any platform, but rather to offer a practical overview of selected AI tools that we have found useful for health professions education research. These tools include Consensus, Elicit, Scite, Litmaps, ResearchRabbit, Paperpal, Jenni AI, Claude, ChatGPT, Gemini, and Microsoft Copilot. We group these tools into 6 functional categories and discuss their strengths and limitations across different stages of the scholarly workflow, drawing on vendor documentation and our own experience rather than formal comparative testing, while emphasizing the importance of selecting tools that align with specific scholarly tasks. Since overreliance on AI risks diminishing creative scholarly competencies, we distinguish productive cognitive support, in which AI reduces routine workload, from epistemic off-loading, in which interpretation and judgment are delegated to AI, and argue for using AI as a scaffold that enhances, rather than replaces, intellectual engagement. We also address some ethical considerations, including hallucinated content, citation inaccuracies, transparency in disclosure, and subscription costs. Ultimately, responsible integration of AI into scholarship will require not only appropriate tool selection but also the continued cultivation of the critical thinking and scientific reasoning that define academic scholarship.

JMIR Med Educ 2026;12:e104203; doi: 10.2196/104203

Keywords: AI; medical education; clinician educators; literature review; academic writing; AI tools; scholarship

Introduction

Academic clinician educators occupy a unique and demanding role at the intersection of clinical practice, teaching, and scholarship, requiring them to balance multiple professional identities simultaneously [1]. They serve as the backbone of the teaching mission in health professions institutions while also being expected to contribute to health professions education research, often with limited formal training in research methodology [2,3]. The barriers are well-documented and concrete. A recent scoping review of author development programs similarly concluded that successful educational scholarship depends not only on individual motivation but also on institutional investment in structured faculty development, mentorship, collaborative writing opportunities, and a culture that values educational scholarship [4]. Clinician educators rarely have protected time to conduct the prerequisite literature search and review that scholarly work requires, and they frequently struggle to find mentoring or editorial help to refine their writing for peer review [2,3,5,6]. The result is a familiar tension: the expectation to produce rigorous, peer-reviewed scholarship under real constraints of time, resources, and methodological support.

AI tools offer a compelling, if still evolving, way to address both of these pressures [7]. Over the past few years, the landscape of literature reviews and academic writing support has changed substantially. Where researchers once navigated static bibliographic indexes by hand, they can now use AI-assisted platforms that rapidly identify, organize, and summarize large bodies of evidence and large language model (LLM) assistants that help create outlines, restructure, and clarify writing. These tools do not replace research expertise, but they have particular utility for 2 stages of the scholarly workflow, early-stage literature exploration, and academic writing, where current AI capabilities are most mature and can meaningfully reduce the logistical burden. Ott et al [8] describe the emergence of AI-mediated writing environments as fundamentally reshaping academic writing practices, emphasizing that educators must intentionally develop pedagogies that preserve critical thinking and reflective writing rather than simply improve efficiency. As these technologies become increasingly integrated into academic practice, their successful adoption also depends on appropriate faculty development. Studies suggest that clinician educators benefit most from structured AI training that combines hands-on experimentation with discussion of ethical reasoning, critical appraisal of AI-generated outputs, and human oversight rather than focusing solely on technical proficiency [9]. This reinforces the need to view AI as a tool that complements scholarly expertise through thoughtful implementation and ongoing professional development.

The gap this viewpoint addresses is one of awareness of the use of AI for scholarly purposes. Many clinician educators do not know about all the tools available and which ones are actually built for scholarly work. Figure 1 provides an overview of the integration of AI tools into the scholarly process. Thus, the aim of this viewpoint is to provide clinician educators with practical guidance for selecting and responsibly integrating AI tools across the scholarly workflow. Specifically, we seek to (1) clarify the capabilities and limitations of commonly used AI platforms, (2) recommend how different tools can be matched to different stages of literature review and academic writing, and (3) highlight the ethical considerations necessary for responsible AI-assisted scholarship. The recommendations presented in this viewpoint are informed by 3 complementary sources: published literature on AI in medical education and scholarly writing, publicly available resources describing the capabilities of AI platforms, and the collective experience of the authors as clinician educators actively engaged in health professions education scholarship.

Tool selection was purposive rather than systematic. Our primary criterion was relevance to a specific stage of the scholarly workflow: we sought representative platforms for each of the 6 functional categories described below (Table 1), so that topic orientation, literature discovery, citation mapping, citation appraisal, and academic writing were each illustrated by at least one example. Within each category, selection was further guided by the authors' direct experience with the platform in their own scholarship, which allowed us to comment on practical use; by accessibility to individual clinician educators through a free tier, free trial, or individual subscription; and by each platform's visibility in academic settings, where these tools are widely used or discussed. Inclusion does not imply that a platform outperforms alternatives in its category, and other platforms may serve the same functions equally well. AI tools were assessed across key dimensions: scholarly function, data sources and coverage, citation capabilities, usability, privacy considerations, cost, and workflow suitability. Information was obtained from vendor websites, published literature, and authors' experience. We did not conduct independent benchmarking or comparative testing; thus, findings should be interpreted as practical guidance rather than formal effectiveness comparisons. Accordingly, descriptions of platform features and data sources are drawn from vendor documentation, whereas observations about how well a tool performs in practice reflect our own experience and are identified as such.

Our aim is not a systematic or exhaustive review of these tools, but practical guidance for the AI-novice clinician educator on how to select, budget for, and responsibly deploy tools suited to specific tasks in health professions education scholarship, with attention to their differing functions and limitations. While AI may reduce the barriers to generating scholarship by accelerating literature discovery, information organization, and manuscript preparation, the scholarly process is fundamentally more than the efficient production of academic outputs. Scholarship requires critical reading, evaluation of evidence, synthesis of diverse perspectives, construction of coherent arguments, and the application of

scientific judgment. Throughout this viewpoint, we argue that AI should be viewed as a tool that allows clinician educators to devote more time to the higher-order cognitive work that defines rigorous academic inquiry. We are a group of health professions educators and researchers and do not claim to have AI expertise; therefore, for the purpose of this viewpoint, we relied on data provided on the various platforms' web pages.

Figure 1. A practical 6-step workflow for integrating AI tools into health professions education scholarship. This figure was developed from the authors' experience and a review of currently available AI platforms.

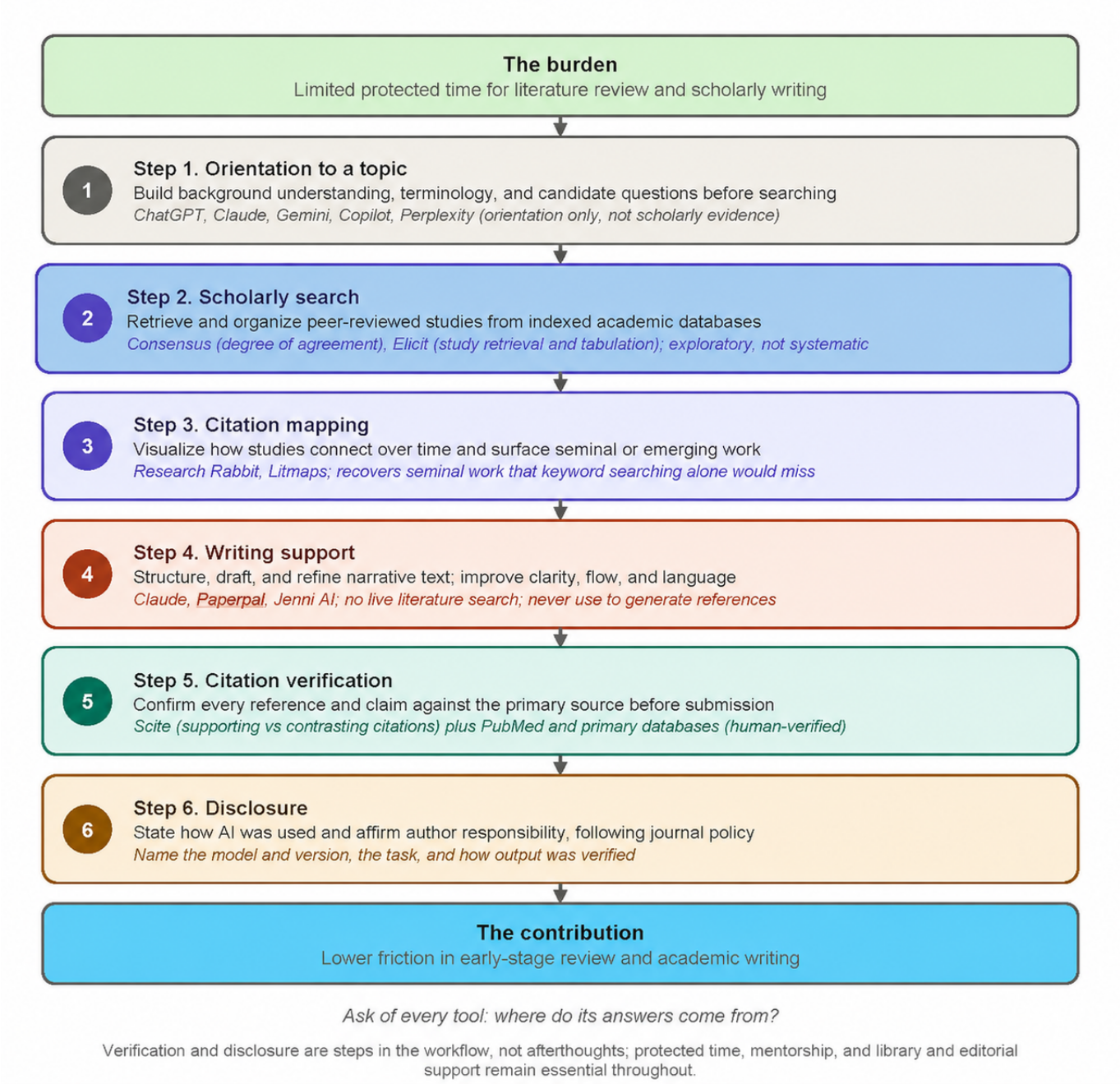


Table 1. Task-oriented guide to selected AI tools for clinician educator scholarship, grouped by functional category.[a]

Tool	Primary purpose	Key strengths	Key limitations	Data source	Indexed scholarly sources[b]	Approximate cost
Scholarly discovery platforms (Figure 1, Step 2: scholarly search)						
Consensus	Gauge scientific agreement on a question and summarize findings across studies	Distills agreement across studies; user-friendly for nonexperts; large scholarly corpus	Best for questions answerable by existing literature; less suited to emerging topics	Semantic Scholar, OpenAlex, and PubMed (roughly >200 million papers)[c]	Semantic Scholar, PubMed, and OpenAlex	Free tier (limited AI analyses); Pro US $15/mo or US $120/y[d]
Elicit	Structured literature search and evidence-table extraction	Draws on curated scholarly corpora; semantic search; structured extraction of study details	Some nuance lost in autosummarization; coverage depends on indexed sources	Semantic Scholar, PubMed, and OpenAlex (>138 million papers; 545,000 trials)[e]	Semantic Scholar, PubMed, and OpenAlex	Free tier (limited extractions); Plus from US $12/mo (~US $10/mo annual)[e]
Citation-context tools (Figure 1, Step 5: citation verification)						
Scite	Citation-context analysis and critical appraisal	Shows whether citing papers support, contrast, or mention a claim; useful for appraisal and reference checking	Not a discovery-first search tool; coverage varies by discipline	Proprietary Smart Citation index built from PubMed, Crossref, Unpaywall, and publisher full text (>1.6 billion citation statements)[f]	PubMed (via Crossref and publisher agreements); not Semantic Scholar or OpenAlex	Free 7-day trial; US $20/mo[f]
Citation-network visualization tools (Figure 1, Step 3: citation mapping)						
Litmaps	Citation-network visualization and literature monitoring over time	Temporal mapping of citation networks; alerts for new papers; helps surface emerging clusters	Visualization tool, not a synthesis tool; requires seed article(s) to build maps	Citation data across >270 million articles (Semantic Scholar, Crossref, and OpenAlex)[g]	Semantic Scholar and OpenAlex[h] (PubMed reached indirectly)	Free tier (2 maps/mo, 100 articles/map); Pro US $10/mo or US $120/y[g]
ResearchRabbit (RR)	Citation-network visualization and discovery	Maps how papers relate through citations and authorship; identifies foundational works	Visualization tool, not a synthesis tool; does not generate summaries	Semantic Scholar/OpenAlex[i]	Semantic Scholar, OpenAlex[h] (PubMed reached indirectly)	Free, unlimited searches (1 project; paid RR+ for more capacity)[i]
Web-grounded search tools (Figure 1, Step 1: orientation only)						
Perplexity	Broad topic overview via general web search (generic, not a dedicated scholarly tool)	Fast, conversational; good for initial orientation to an unfamiliar topic	Defaults to open-web search and may include non–peer-reviewed content; citation reliability varies; advise against as primary tool for a robust literature search	Open web by default; optional academic mode draws on Semantic Scholar[j]	None[k] by default; academic mode uses Semantic Scholar only	Free tier; Pro US $20/mo or US $200/y[j]
General-purpose LLM[l] assistants (Figure 1, Steps 1 and 4: orientation and writing support)						
Claude (Anthropic)	Academic writing support, synthesis, and paraphrasing	Strong natural-language generation; useful for drafting, editing, and structuring text; accepts uploaded documents	No live literature search; may fabricate citations; all outputs require human verification	Model training data plus user-provided documents (no live database integration)[m]	None[k] (training data and uploaded documents only)	Free tier; Pro US $20/mo (discounted annual rate available)[m]
Gemini (Google)	General-purpose writing and synthesis assistant	Strong at ingesting and synthesizing large-source volumes; grounds answers in current web results via Google Search	Original prose generally rated below Claude and ChatGPT; web grounding is not a curated scholarly corpus; may fabricate citations	Google Search grounding (open web; no curated scholarly database)[n]	None[k] (Google Search/open web)	Free tier; Google AI Pro US $19.99/mo (AI Plus US $7.99/mo)[n]
Microsoft Copilot	General-purpose writing assistant integrated with Microsoft 365	Drafts and edits within Word, Excel, Outlook, and Windows; useful where work already lives in Microsoft 365	Not a scholarly search tool; web answers may include non–peer-reviewed content; may fabricate citations; outputs require verification	OpenAI GPT model family plus open web via Bing (no curated scholarly database)[o]	None[k] (Bing/open web)	Free tier; Copilot Pro US $20/mo[o]

(page number not for citation purposes)

Tool	Primary purpose	Key strengths	Key limitations	Data source	Indexed scholarly sources[b]	Approximate cost
Academic writing assistants (Figure 1, Step 4: writing support)						
Jenni AI	Purpose-built academic writing assistant with inline citation generation	Academic-tuned autocomplete and paraphrasing; inserts citations from the user's uploaded research library; exports to Word and LaTeX	Cites only from the user's library, not a live external search; citations require verification	User-uploaded research library plus journal and web citation sources; runs on third-party models (eg, ChatGPT [GPT-4], Gemini)[p]	None of the 3 directly[k]; cites from the user's uploaded library and journal/web sources	Free tier (200 AI words/day); unlimited US $20/mo or US $12/mo annual (US $144/y)[p]

[a]Platforms are grouped into the 6 functional categories described in the text, each labeled with the corresponding step of the Figure 1 workflow; they were selected purposively as illustrative examples, and neither their inclusion nor their order implies a ranking. Entries are based on the authors' experience and publicly available platform documentation (June 2026); the "Key strengths" and "Key limitations" columns are descriptive and do not reflect independent comparative testing. The table is intended as a practical guide for clinician educators and should not be interpreted as a systematic review, comprehensive market comparison, formal performance evaluation, or endorsement of AI tools. Information is current as of June 2026. All pricing and data source details are subject to change; users should consult each platform's documentation for the most current information.

[b]The "Indexed scholarly sources" indicates which of the 3 major scholarly indexes (Semantic Scholar, PubMed, and OpenAlex) are used by each platform.

[c]Consensus searches a database of roughly 200 million peer-reviewed papers drawn from Semantic Scholar, OpenAlex, and PubMed (over 37 million PubMed records are searchable within Consensus), supplemented by its own web crawl [10].

[d]Consensus pricing: a permanent free tier provides basic search but meters the AI-powered analyses (Pro and Deep Searches); Pro is US $15/mo or US $120/y [10].

[e]Elicit searches >138 million papers and ~545,000 clinical trials across Semantic Scholar, PubMed, and OpenAlex. The free Basic tier allows unlimited searches and summaries but caps structured extractions at about 20/mo; Plus starts at US $12/mo, reduced with annual billing [11].

[f]Scite uses a proprietary Smart Citation index of >1.2 billion citation statements. There is no permanent free tier; after a 7-d free trial, the individual plan is US $20/mo, reduced to about US $12/mo (US $144/y) with annual billing [12].

[g]Litmaps draws on citation data across >270 million articles. The free tier is limited (basic search with up to 20 inputs and 2 maps per month with 100 articles per map); Pro is US $10/mo or US $120/y [13].

[h]PubMed coverage is often obtained indirectly through Semantic Scholar or OpenAlex rather than via direct searches of PubMed.

[i]ResearchRabbit visualizes citation and authorship networks and has long been free with unlimited searches; its pricing page now lists a free plan (unlimited searches, one project) alongside a paid RR+ tier for additional capacity [14].

[j]Perplexity is a general-purpose answer engine that searches the open web by default; an optional academic focus mode restricts results to peer-reviewed sources indexed in Semantic Scholar. There is a free tier, with Pro priced at US $20/mo or US $200/y [15].

[k]"None" indicates a tool that does not search a curated scholarly index and instead relies on the open-web content, model training data, or a user-supplied library.

[l]LLM: large language model.

[m]Claude does not perform live literature searches and may produce inaccurate citations; there is a free tier, with Pro at US $20/mo and a discounted annual rate available [16].

[n]Gemini is a general-purpose model that grounds answers in Google Search results (the open web) rather than a curated scholarly corpus; it offers a free tier, with Google AI Pro at US $19.99/mo and a lower-cost Google AI Plus tier at US $7.99/mo [17].

[o]Microsoft Copilot is a general-purpose assistant powered by OpenAI ChatGPT models, integrated across Microsoft 365 and drawing on the open web via Bing rather than a scholarly database; it offers a free tier, with Copilot Pro priced at $20/mo [18].

[p]Jenni AI generates citations only from the user's uploaded research library rather than from a live external database, and runs on third-party models. The free tier is limited (about 200 AI-generated words/day, 10 stored documents, and 10 AI-chat messages); the Unlimited plan is US $20/mo billed monthly or US $12/mo (US $144/y) billed annually [19]

Understanding the AI Landscape for Scholarship: Capabilities and Limitations of Common AI Platforms

Overview

A foundational issue when evaluating AI tools for scholarship is recognizing that they differ substantially in how they retrieve and process information. This distinction matters for academic use, yet it is frequently overlooked. The single most useful question a clinician educator can ask of any tool is: "where do its answers come from?"

To make these differences explicit, we group the platforms discussed in this viewpoint into 6 functional categories, defined by the primary task each tool performs and by where its outputs come from (Table 1). The categories are ordered from tools anchored most directly in indexed scholarly literature to those that primarily generate text, and Figure 1 shows where these tools fit in the scholarly workflow. Importantly, these categories describe *primary* functions rather than fixed boundaries: Microsoft Copilot, for example, serves both as a general-purpose assistant and as a web search interface, and general-purpose assistants increasingly add web search while scholarly platforms add generative summaries. We therefore use "AI tools" as an umbrella term only when a point applies across categories, and otherwise refer to the specific category.

The alignment of an AI platform and its overall function are critical for appropriate use in a scholarly setting. One challenge with using AI platforms for scholarly purposes is that general-purpose LLM assistants (eg, Gemini; Google) and web-grounded search tools (eg, Perplexity) could be easily mistaken for research tools. While these tools are good for generating summaries and answering questions, these tools are generic [15,17] and not natively anchored to curated scholarly databases. Therefore, their outputs may include non–peer-reviewed web content. In contrast, authors should consider AI tools and platforms that draw on indexed academic corpora or citation networks [10-14], which we recommend to clinician educators based on our group's research experience. These include scholarly discovery platforms (eg, Elicit and Consensus), citation-context tools (eg, Scite), and citation-network visualization tools (eg, Litmaps and ResearchRabbit). Finally, academic writing assistants (eg, Paperpal and Jenni AI) are built specifically to refine academic language and provide citation-related support within manuscript drafts [19,20].

This heterogeneity means clinician educators should approach tool selection with clarity of purpose. A tool well-suited to mapping a citation network (Litmaps) serves a different function than one suited to generating a narrative description (Claude; Anthropic) or to appraising the credibility of individual citations (Scite). A task-oriented summary of these platforms, grouped by functional category and including their primary data sources, strengths, limitations, and approximate costs, is provided in Table 1; it is intended to help match tools to tasks rather than to rank platforms against one another. Guidance on how to integrate the various platforms based on their functional categories is included in the following sections.

Using AI Tools for the Scholarly Workflow: Early Literature Exploration

For clinician educators beginning a new project, a staged approach to AI-assisted literature review is both practical and appropriate. The goal of an early-stage review is typically to orient oneself to a topic, identify key themes and gaps in the literature, and locate seminal studies. It is important to distinguish this exploratory phase from formal evidence synthesis. AI tools can efficiently support the early stages: topic orientation, literature discovery, citation mapping, and organization of the literature. However, they are not substitutes for rigorous, protocol-driven evidence syntheses such as systematic reviews or scoping reviews, which require predefined research questions, comprehensive search strategies, explicit inclusion and exclusion criteria, duplicate screening, critical appraisal, and transparent reporting according to established methodological frameworks. For formal evidence syntheses, collaboration with an information specialist or librarian can be particularly valuable in developing comprehensive search strategies, selecting appropriate databases, and ensuring that searches are reproducible and appropriately documented.

Selecting the appropriate AI platform for its purpose can help streamline this process, including choosing platforms that curate scholarly corpora. Elicit, a scholarly discovery tool, draws on more than 138 million academic papers and roughly 545,000 clinical trials from sources including Semantic Scholar, PubMed, and OpenAI, so its outputs are anchored in indexed academic literature [11]. Consensus likewise searches a large scholarly corpus, indexing roughly 200 million papers through Semantic Scholar, and is designed to distill evidence across studies and gauge the degree of scientific agreement on a question [10]. Scite, a citation-context tool, analyzes how a paper has been cited, classifying citing statements as supporting, contrasting, or mentioning, which makes it valuable for critical appraisal rather than discovery [12], in contrast with Elicit and Consensus. Litmaps and ResearchRabbit are citation visualization platforms that map citation networks and reveal how studies connect over time [13,14]. Neither is primarily an AI synthesis tool in the LLM sense; their strength is surfacing seminal or emerging work that keyword searches alone might miss.

By contrast, Perplexity and Microsoft Copilot, when used for web queries, operate primarily as web search tools [15, 18]. By default, they search the open web and offer a fast, conversational interface, although their results may include non–peer-reviewed content. Perplexity does offer an optional academic focus mode that restricts results to peer-reviewed sources indexed in Semantic Scholar, but, in its default configuration, it is not anchored to a curated scholarly corpus in the way Elicit or Consensus are. In our experience, Perplexity is useful for initial orientation on an unfamiliar topic, but unless "academic mode" is deliberately enabled its outputs should be treated as a starting point, not as a reliable

source of scholarly citations. General-purpose LLM assistants such as Claude, ChatGPT (Open AI), and Gemini, in turn, are designed primarily to generate and refine narrative text rather than to search curated scholarly databases. Because these general-purpose LLM assistants are not primarily designed for literature searches, it is important to recognize that they generate text from their training data, from any documents the user provides, and, where web search is enabled, from open-web results rather than a curated scholarly index. Caution should be exercised when reviewing recommended citations, as LLMs are known to produce plausible sounding but inaccurate references or hallucinations [21]. All AI-generated citations must be verified against primary databases before inclusion in a manuscript.

Using AI Tools for the Scholarly Workflow: Academic Writing

Beyond literature synthesis, AI writing tools fall into 2 broad groups that serve different purposes, and clinician educators benefit from understanding the distinction. The first group comprises general-purpose LLM assistants such as Claude [16], ChatGPT [22], and Gemini [17]. These are sometimes called frontier models, meaning the most capable general-purpose AI systems available at a given time, built by the leading developers and updated frequently. Their strengths for writing lie in improving prose clarity, restructuring arguments, and paraphrasing. We have not compared these assistants formally, and the observations that follow reflect our own use rather than independent evidence. In our experience, Claude has been helpful for following detailed instructions across multiple uploaded documents and synthesizing them succinctly, ChatGPT for a wide range of general tasks, and Gemini for synthesizing large volumes of source material and grounding answers in current web results. Relative performance varies with the task, the prompt, and the model version, and frequent updates can quickly change how these tools compare; these impressions are therefore best treated as starting points for readers' own evaluation rather than as a ranking.

The second group comprises tools built specifically for academic writing. According to its developer, Paperpal is trained on published, peer-reviewed articles rather than general web text and is designed to handle academic writing, discipline-specific terminology, and submission-readiness editing more reliably than general-purpose models, and it integrates directly into Microsoft Word and Google Docs [20]. These are the vendor's descriptions; we did not independently evaluate Paperpal's performance against that of general-purpose models. Jenni AI occupies a related niche: it pairs drafting assistance with in-line citation generation from academic sources, an approach intended to reduce the fabricated-citation problem that affects general-purpose models; however, its citations still require manual verification [19]. One pattern we have found practical is to draft and structure with a strong general-purpose model and then refine academic language with a purpose-built tool, rather than expecting any single tool to do both well; responsibility for academic rigor remains with the author.

A practical consideration for getting the most from these tools is the quality of the input. As AI becomes increasingly integrated into health care and academic work, prompt engineering may be considered an important professional competency rather than simply a technical skill [23]. Meskó [24] argues that the ability to design clear, structured prompts is becoming essential for health care professionals to effectively and responsibly interact with LLMs. A structured, specific prompt that clearly states the desired role, task, context, and output format consistently produces more useful results than a vague request [24]. Specifying the target audience (eg, medical educators unfamiliar with a methodology), the desired tone (scholarly but accessible), and the output format (a 250-word narrative paragraph with American Medical Association citations) will yield more actionable text than a general request to "summarize this topic." AI tool development is moving rapidly toward agentic workflows, systems that can autonomously complete multistep tasks [25]. While careful prompting remains an important competency for current LLM interfaces, clinician educators should anticipate a transition toward collaborative human-AI workflows in which prompting becomes one component of interacting with increasingly sophisticated AI agents [26].

Using AI Tools for the Scholarly Workflow: A Practical Example and Budgeting

Here is a practical example of a reasonable workflow and task-tool matching: consider a clinician educator developing a project on whether structured feedback improves resident learning. The first step may be topic orientation, using tools such as Perplexity or Consensus to clarify key concepts and map the general landscape of evidence. For literature discovery and organization, Elicit can help identify relevant studies and extract key characteristics. Once core papers are identified, ResearchRabbit or Litmaps can map citation networks and highlight influential or emerging work. Scite can then be used to assess how studies are cited in the literature, including whether findings are supported, challenged, or simply referenced. For academic writing support, general-purpose LLM tools such as Claude, ChatGPT, or Gemini can then help summarize and organize information retrieved during early literature exploration, assist in drafting narrative text, and improve clarity and flow. Tools like Paperpal or Jenni AI may help refine academic language and provide citation-related support. The platforms named in this example illustrate each functional category; other tools in the same category can often be substituted.

When combining platforms, budget matters. For example, a clinician educator using these tools over 1 month of active literature work might pair Elicit Plus (about US $12/month) with Scite (US $20/month) and a Claude Pro subscription (US $20/month), for a total monthly cost of US $52, while they may use the free tiers of Consensus, Litmaps, and ResearchRabbit. Many institutions do not subsidize these costs; therefore, clinician educators should factor platform expenses into project planning. Current pricing for each platform is summarized in Table 1.

Preserving Scholarly Judgment and Responsible AI Use

The integration of AI into scholarly work raises ethical questions that clinician educators must deliberately consider. The risks are real: generative AI tools can introduce hallucinated content, produce inaccurate or fabricated citations, and reflect biases present in their training data. Outputs may misrepresent the state of evidence, particularly where the literature is sparse or contested. A systematic review by Basu and Huynh [21] concluded that effective mitigation requires both technical safeguards and active human oversight, underscoring that AI should augment rather than replace scholarly judgment. Cross-checking all AI-generated content against primary sources is a baseline requirement for scholarly integrity. Verification is more than a procedural step; it is an epistemic competency. Responsible AI-assisted scholarship requires evaluating whether claims are evidence-based, distinguishing scientific evidence from the fluency or authority of the model, and recognizing uncertainty, gaps, and alternative interpretations in the primary literature. AI-generated outputs may be coherent and confident even when evidence is incomplete or inaccurate. Critical engagement therefore extends beyond asking "Is this accurate?" to asking "What supports this claim?," "How strong is the evidence?," and "What has been omitted or oversimplified?" As described in recent work by Rind et al [27], the continuum of patterns of AI use in higher education ranges from blind trust to structured claim-evidence evaluation. The most mature users inhabit the role of critical evaluator, applying academic reasoning and discipline to their engagement with AI tools [27]. This highlights the importance of these skills in health professions education and AI-mediated scholarship.

Transparency in disclosure is equally important, and the norms are actively evolving. A recent survey of more than 5000 researchers conducted by *Nature* found respondents divided on what constitutes appropriate AI use, with views ranging from permitting grammar and clarity assistance to prohibiting the use of AI for drafting any substantive text [28]. Clinician educators should follow the disclosure policies of their target journals and institutions and should clearly acknowledge AI tool use in their manuscripts, as we have done here. In addition, clinician educators must exercise caution when using AI platforms with unpublished manuscripts, learner submissions, institutional documents, or other potentially sensitive information. Unless institutional policies explicitly permit such use, confidential or identifiable content should be deidentified or excluded from AI platforms, and users should ensure compliance with institutional data governance, privacy regulations, and publisher requirements.

Mehta et al [29] proposed an AI competency framework for physicians that emphasizes critical appraisal of AI outputs, ethical reasoning, interdisciplinary collaboration, and responsible deployment of AI technologies, arguing that these skills are essential for future health care professionals. The tools described in this viewpoint are a means, not ends; their value depends entirely on the judgment of the humans using them. Rather than simply reducing the workload, AI can redistribute the cognitive work involved in scholarship. This can free clinician educators to devote more attention to intellectually demanding tasks, such as interpreting evidence, synthesizing competing perspectives, or constructing scholarly arguments. Excessive reliance on AI risks shifting researchers from active engagement with the literature to passive acceptance of machine-generated summaries [30]. Over time, this may diminish the very competencies that academic scholarship is intended to cultivate, including critical appraisal, scientific reasoning, and independent judgment.

These 2 possibilities can be framed more precisely as a distinction between productive cognitive support and epistemic off-loading. Off-loading cognitive work is not inherently harmful: in their guidance on supervising learners' use of AI, Abdulnour et al [31] note that it can ease demands on working memory and free attention for more demanding tasks, whereas off-loading complex reasoning risks automation bias, deskilling, never-skilling, and misskilling. Productive cognitive support delegates routine, well-defined, and verifiable work, such as deduplicating search results, extracting predefined study characteristics, reformatting references, or tightening prose, while the scholar remains responsible for deciding what the evidence means. Epistemic off-loading delegates the interpretive and evaluative work that constitutes scholarship itself: judging which studies are relevant and credible, weighing the strength of evidence, recognizing what is missing or contested, and deciding which conclusions are warranted. The distinction lies less in the tool than in the task delegated and the user's stance toward the output. Asking an AI tool to summarize a set of papers is productive support when the researcher has read the key studies and uses the summary to organize their own analysis; the same request becomes epistemic off-loading when the summary substitutes for reading and its conclusions pass unexamined into a manuscript.

Epistemic off-loading is difficult to recognize from the inside. When AI tools are trusted to survey and summarize the literature, researchers become vulnerable to illusions of understanding, believing they understand more than they do [32]. The continuum described by Rind et al [27] maps onto this distinction: at the blind-trust end, interpretation and judgment are effectively off-loaded, whereas at the critical-evaluation end, the researcher retains epistemic authority and treats each output as a set of claims to be tested against primary evidence. As a practical default, we suggest that clinician educators reserve for themselves the framing of the research question, the appraisal of key studies, the weighing of conflicting findings, and the formulation of conclusions, using AI for retrieval, organization, and language refinement around those decisions.

Institutions also play a critical role in the adoption of AI tools for health professions education research. We call for academic medical centers to build interdisciplinary structures that bring together faculty, librarians, information technology staff, and administrators to develop policies on AI

use, ensure access to AI platforms, and provide training that enables critical, informed engagement. Cole maintains that AI changes—not replaces—the educator's role, and that institutions must rethink faculty development and scholarly work [1]. Access to AI is most likely to improve scholarship when embedded within institutions that also invest in faculty development, research collaboration, and robust academic support systems. In this context, AI becomes an amplifier of scholarly capacity rather than a substitute for institutional responsibility. Without that institutional scaffolding, individual clinician educators are left to navigate a rapidly changing landscape largely on their own. Access is also not equitable: subscription costs, while modest by some standards, represent a meaningful barrier for clinician educators at underresourced institutions or in lower-income settings. Negotiated institutional access is an important equity consideration for the medical education community to address. Accordingly, strategies for integrating AI into clinician educator scholarship should remain flexible enough to accommodate the differing resources, infrastructure, and support systems available across institutions.

AI Disclosure

How to disclose AI use is itself a source of uncertainty, and clear guidance helps clinician educators disclose with confidence. A recent AMEE (International Association for Health Professions Education) Guide sets out a framework of good practice that specifies the elements a complete disclosure should contain: name the AI model and version used, describe the specific activities for which it was used, declare consent if any participant data were uploaded, describe the data-protection measures taken, state when AI was deliberately not used in cases where a reader might assume otherwise, reference the relevant journal or institutional policies, include a final statement of author responsibility, and retain a record of prompts and interactions in case editors request verification [33]. Textbox 1 provides an adaptation with a fill-in template, so that clinician educators can work through it point by point.

Textbox 1. Practical framework for transparent disclosure of AI use in medical education scholarship. Adapted from the AMEE (International Association for Health Professions Education) Guide on AI disclosure in academic publishing, this textbox provides a structured checklist and fill-in template to assist clinician educators in documenting AI use during scholarly work [33].

1. Specify the AI model used. Clearly name the AI system, including version numbers or details of any custom or local model.
2. Describe the specific activities for which AI was used. Give precise detail of how the AI contributed, and avoid vague descriptions.
3. Declare consent if participant data were uploaded. Confirm that appropriate consent was obtained for uploading any learner, patient, or participant data.
4. Describe data-protection measures taken. Explain safeguards such as disabling model-training settings or using privacy-preserving mechanisms.
5. State when AI was not used, where relevant. Clarify omissions in cases where a reader might reasonably assume AI involvement.
6. Reference relevant journal, publisher, or institutional AI policies. Affirm that the AI use aligns with applicable guidelines.
7. Include a final statement of responsibility. Assert full responsibility for the accuracy, integrity, and validity of the content.
8. Maintain a record of prompts and interactions. Keep documentation of AI interactions in case editors request verification.

AI Use Disclosure Fillable Template

1. AI tool used: ……
2. Specific activities for which AI was used: ……
3. Consent obtained for uploaded data (if applicable): ……
4. Data-protection measures taken (if applicable): ……
5. Statement on what AI was not used for (optional): ……
6. Statement of compliance with journal or institutional AI policies: ……
7. Final responsibility statement: ……
8. Location of prompts or records (if retained): ……

A study of editors of health professions education journals offers practical, editor-derived guidance on how much to disclose and where [34]. Three principles recur: name the tool, specify the task, and attest responsibility. Beyond those basics, the study highlights 2 judgments that authors find difficult. The first is necessity, or whether a given use needs to be disclosed at all. Routine spelling and grammar assistance is generally not seen as warranting disclosure by these editors, but any use that shapes ideas, the structure of an argument, or scholarly content does; when the distinction is unclear, the editors' advice is simply to disclose. The second is sufficiency, or how much detail to give. Here the guidance is to scale the detail to the nature of the use: a brief note suffices for technical help such as shortening an abstract, whereas more intellectual uses, such as drafting a section of the discussion or generating counterarguments, call for a fuller account of what was prompted and how the output was checked.

Several further points make a disclosure more credible. It helps to indicate not only what AI produced but also how its output was verified, since verification is a process rather than a promise. For example, this can include confirming AI-suggested claims against a primary database and editing where the model oversimplified. Editors generally prefer the disclosure to appear in the Methods section or another logical place in the manuscript, rather than buried in the acknowledgments, and they prefer that any checkbox or text entered during online submission be consistent with the disclosure in the manuscript itself. Finally, the editors in this study emphasized that disclosure is meant to be nonpunitive; norms are still being worked out collaboratively between authors and journals, and disclosing thoughtfully helps the field develop shared standards rather than exposing the author to penalties [34].

Counterarguments and Limitations of This Viewpoint

We acknowledge that our perspective is experience-informed rather than empirically derived. We have not conducted a systematic review of these tools, nor have we collected data from clinician educators on their experiences. Our characterizations reflect information obtained directly from the vendor sites and our own best understanding at the time of writing, but in a rapidly evolving field, specific features, data sources, and pricing change quickly, and some characterizations may already have shifted. Accordingly, the tools included should be viewed as illustrative examples rather than a comprehensive inventory of available AI platforms, and their inclusion should not be interpreted as an endorsement of or ranking over alternative tools. Likewise, our observations about the relative strengths of particular platforms come from our own use rather than head-to-head testing, and because selection drew partly on our experience, other capable platforms may have been omitted. Formal comparative evaluations using predefined scholarly tasks and outcome measures would be needed to support claims that one platform performs better than another.

A legitimate counterargument is that these tools may lower the barrier to producing low-quality work at high volume, contributing to publication numbers rather than genuine knowledge advancement, and that reliance on AI for literature synthesis may encourage superficial engagement with primary sources [35]. These are valid cautions. We do not advocate AI as a shortcut; we see it as a scaffold that is most valuable when used by clinician educators who remain actively engaged in critically reading, interpreting, and contextualizing the literature.

Conclusions

AI tools offer genuine potential to support clinician educators in health professions education scholarship, particularly in the early stages of literature review and in academic writing. Realizing that potential requires selectivity: understanding what each tool actually does, what data it draws from, and where its limitations lie; distinguishing purpose-built scholarly tools from generic web-search assistants; and matching each category of tool to the task at hand. It also requires intellectual honesty about the risks of uncritical use, and, in particular, a deliberate distinction between delegating routine work to AI and delegating the interpretation and judgment that constitute scholarship. We encourage clinician educators to experiment with these platforms thoughtfully, to maintain rigorous verification practices, and to advocate within their institutions for the governance structures and training that responsible AI integration demands. Importantly, as AI capabilities continue to evolve, the future of clinician educator scholarship will depend on using AI to augment, rather than replace, critical reading, methodological rigor, evidence synthesis, and the human judgment that remains central to high-quality research. Future work should explore how clinician educators at varying career stages and institutional contexts adopt and adapt these tools, and what forms of institutional support most effectively enable ethical and productive use.

Generative AI Use

The authors used Claude (Anthropic) to assist with drafting and editing portions of this manuscript and Microsoft Copilot to assist with title development and information gathering for comparing tools (accessed March-May 2026). These tools were used to generate initial ideas, help structure an outline, and provide language and grammar suggestions to refine sections of the text. All conceptualization, interpretation, and argumentation were developed by the authors. The authors critically reviewed, revised, and verified all AI-assisted content, including checking factual claims against source materials to ensure accuracy, originality, and coherence with journal guidelines. The authors take full responsibility for the final content and interpretations. AI is a rapidly evolving field; the tool features, pricing, and database integrations described here may change after the date of publication.

Funding

The authors declared no financial support was received for this work.

Conflicts of Interest

None declared.

References

1. Cole R. Is AI replacing faculty? Rethinking faculty roles in medical education. Med Educ. Sep 2026;60(9):955-957. [doi: 10.1111/medu.70220] [Medline: 41934214]
2. Harvey Bluemel A, Burton OE, Burford B, et al. Barriers and facilitators to establishing a clinical academic career in clinical education research in the UK: a focus group study. Med Teach. Oct 2024;46(10):1369-1377. [doi: 10.1080/0142159X.2024.2308783] [Medline: 38359431]
3. Chen YS, Liao KC, Yau SY. From hesitation to participation: a narrative review of facilitators and barriers for healthcare professionals' engagement in medical education research. BMC Med Educ. Feb 25, 2025;25(1):302. [doi: 10.1186/s12909-025-06792-w] [Medline: 40001074]
4. Cameron MW, Crowther LN, Huang GC. Faculty development and infrastructure to support educational scholarship: a scoping review on author development. Acad Med. Jan 1, 2023;98(1):112-122. [doi: 10.1097/ACM.0000000000004896] [Medline: 35921165]
5. Reader S, Fornari A, Simon S, Townsend J. Promoting faculty scholarship - an evaluation of a program for busy clinician-educators. Can Med Educ J. 2015;6(1):e43-e60. [Medline: 26451230]
6. Yarris LM, Juve AM, Artino AR Jr, et al. Expertise, time, money, mentoring, and reward: systemic barriers that limit education researcher productivity-proceedings from the AAMC GEA workshop. J Grad Med Educ. Sep 2014;6(3):430-436. [doi: 10.4300/JGME-D-14-00340.1] [Medline: 26279767]
7. Rodman A, Mark NM, Artino AR Jr, Lessing JN. Using generative artificial intelligence in medical education. Acad Med. Feb 1, 2025;100(2):250. [doi: 10.1097/ACM.0000000000005937] [Medline: 39622003]
8. Ott M, Veselka M, Cristancho S, Hibbert K, Lingard L. "Maybe that's the landscape that's coming?": the emergence of writing pedagogies in AI-mediated learning environments. Language and Literacy. May 19, 2026;28(1):76-103. [doi: 10.20360/langandlit29798]
9. Anand R, Bowers N, Manyama MF. Design and evaluation of a faculty development workshop series on integrating generative artificial intelligence in medical education: mixed methods pilot study. JMIR Med Educ. May 6, 2026;12:e89815. [doi: 10.2196/89815] [Medline: 42090298]
10. Subscription plans. Consensus. URL: https://help.consensus.app/en/articles/10087865-subscription-plans [Accessed 2026-06-01]
11. Elicit's source for papers. Elicit. URL: https://support.elicit.com/en/articles/553025 [Accessed 2026-06-01]
12. Scite. URL: https://scite.ai [Accessed 2026-06-01]
13. Litmaps pricing. Litmaps. URL: https://www.litmaps.com/pricing [Accessed 2026-06-01]
14. ResearchRabbit. URL: https://www.researchrabbit.ai [Accessed 2026-06-01]
15. Perplexity. URL: https://www.perplexity.ai [Accessed 2026-06-01]
16. Claude. URL: https://www.anthropic.com/claude [Accessed 2026-06-01]
17. Gemini. URL: https://gemini.google.com [Accessed 2026-06-01]
18. Microsoft Copilot. URL: https://copilot.microsoft.com [Accessed 2026-06-01]
19. Jenni. URL: https://jenni.ai [Accessed 2026-06-01]
20. Durgumahanthi S. Top 5 AI tools for academic writing in 2026. Paperpal. URL: https://paperpal.com/blog/news-updates/top-5-ai-tools-for-academic-writing [Accessed 2026-06-01]
21. Basu S, Huynh B. Mitigating hallucinations in healthcare AI: a systematic review of evidence-based strategies. BMC Health Serv Res. Jun 6, 2026;26(1):1115. [doi: 10.1186/s12913-026-14851-1] [Medline: 42251377]
22. ChatGPT. URL: https://chatgpt.com [Accessed 2026-06-01]
23. Taylor Gonzalez DJ, Djulbegovic MB, Bair H. We need to add prompt engineering education to optimize generative artificial intelligence in medicine. Acad Med. Oct 1, 2024;99(10):1050-1051. [doi: 10.1097/ACM.0000000000005803] [Medline: 38967978]
24. Meskó B. Prompt engineering as an important emerging skill for medical professionals: tutorial. J Med Internet Res. Oct 4, 2023;25:e50638. [doi: 10.2196/50638] [Medline: 37792434]
25. Chow M, Ng O. Beyond chatbots: moving toward multistep modular AI agents in medical education. JMIR Med Educ. Oct 2, 2025;11:e76661. [doi: 10.2196/76661] [Medline: 41037756]
26. Tran N, Patel D, Cooper AB. Questioning the value of prompt engineering in medical education. Acad Med. Jul 1, 2025;100(7):761. [doi: 10.1097/ACM.0000000000006030] [Medline: 40080560]
27. Rind IA, Bhatti J, Chellappan K. From blind trust to critical inquiry: epistemic beliefs and student engagement with ChatGPT in higher education. Jurnal Kejuruteraan. 2026;38(1):01-14. [doi: 10.17576/jkukm-2026-38(1)-01]
28. Kwon D. Is it OK for AI to write science papers? Nature survey shows researchers are split. Nature. May 2025;641(8063):574-578. [doi: 10.1038/d41586-025-01463-8] [Medline: 40369145]

29. Mehta N, Mehta S, Rubenstein A, Wood SK. Not replaced, but reinvented: AI education pathways to prepare future physicians to lead healthcare transformation. Perspect Med Educ. 2025;14(1):849-859. [doi: 10.5334/pme.2233] [Medline: 41278578]
30. Hough J, Culley N, Erganian C, Alahdab F. Potential risks of GenAI on medical education. BMJ Evid Based Med. Dec 1, 2025;30(6):406-408. [doi: 10.1136/bmjebm-2025-114339] [Medline: 41326281]
31. Abdulnour REE, Gin B, Boscardin CK. Educational strategies for clinical supervision of artificial intelligence use. N Engl J Med. Aug 21, 2025;393(8):786-797. [doi: 10.1056/NEJMra2503232] [Medline: 40834302]
32. Messeri L, Crockett MJ. Artificial intelligence and illusions of understanding in scientific research. Nature. Mar 2024;627(8002):49-58. [doi: 10.1038/s41586-024-07146-0] [Medline: 38448693]
33. Cleland J, Driessen E, Masters K, Lingard L, Maggio LA. When and how to disclose AI use in academic publishing: AMEE Guide No.192. Med Teach. Apr 2026;48(4):542-553. [doi: 10.1080/0142159X.2025.2607513] [Medline: 41467560]
34. Lingard L, Driessen E, Oswald K. The blurred thresholds of AI-use disclosure: health professions education journal editors' expectations of necessity and sufficiency. Perspect Med Educ. 2025;14(1):1003-1012. [doi: 10.5334/pme.2326] [Medline: 41424913]
35. Masters K. Artificial intelligence and the death of the academic author. Med Teach. Oct 2025;47(10):1563-1565. [doi: 10.1080/0142159X.2025.2523463] [Medline: 40693529]

Abbreviations

AMEE: International Association for Health Professions Education
LLM: large language model

Edited by Stefano Brini; peer-reviewed by Abbas Husain, Chibuikem Iheagwaram, Irfan Ahmed Rind, Nitin Seam; submitted 09.Jun.2026; final revised version received 13.Oct.2026; accepted 14.Oct.2026; published 24.Nov.2026

Please cite as:
Zaidi Z, Ganjoo R, Brown K, Izzi F, Rubin A, Koutroulis I
Augmented With AI: Practical Guide for Clinician Educators Engaged in Health Professions Education Scholarship
JMIR Med Educ 2026;12:e104203
URL: https://mededu.jmir.org/2026/1/e104203
doi: 10.2196/104203

© Zareen Zaidi, Rohini Ganjoo, Kirsten Brown, Farida Izzi, Anna Rubin, Ioannis Koutroulis. Originally published in JMIR Medical Education (https://mededu.jmir.org), 24.Nov.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Medical Education, is properly cited. The complete bibliographic information, a link to the original publication on https://mededu.jmir.org/, as well as this copyright and license information must be included.