Review | Open Access

Artificial Intelligence Surgery

Valencia-Coronel *et al. Art Int Surg.* 2026;6:36-60 **DOI:**10.20517/ais.2025.112

Artificial intelligence-based tools for optimizing surgical research publications

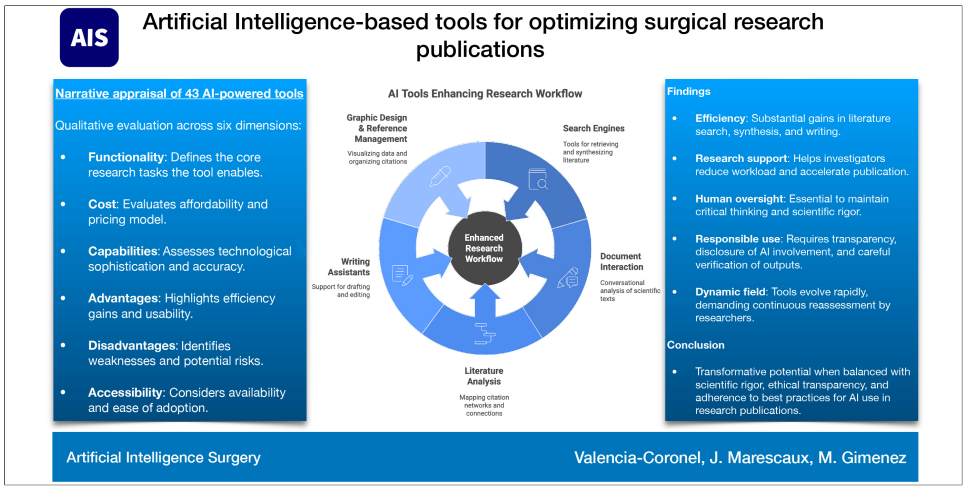
Brandon Valencia-Coronel[1], **Jacques F. Marescaux**[2], **Mariano Gimenez**[2,3,4]

Keywords:
Artificial intelligence and surgery, large language models, surgical research, research tools, scientific writing, research methodology, research efficiency, evidence-based medicine

Citation: Valencia-Coronel B, Marescaux JF, Gimenez M. Artificial intelligence-based tools for optimizing surgical research publications. *Art Int Surg.* 2026;6:36-60. https://dx.doi.org/10.20517/ais.2025.112

Received: 13 Nov 2025 **First Decision:** 19 Dec 2025 **Revised:** 24 Dec 2025 **Accepted:** 12 Jan 2026 **Published:** 16 Jan 2026

Academic Editor: Gaya Spolverato
Copy Editor: Xing-Yue Zhang
Production Editor: Xing-Yue Zhang



Abstract

To evaluate artificial intelligence (AI)-powered tools for optimizing surgical research workflows and establish an evidence-based framework for appropriate AI technology selection throughout the research pipeline. We conducted a structured, qualitative narrative appraisal of 43 AI-powered tools (October 2024-March 2025), categorizing them across five functional domains: (1) scientific search engines, (2) document interaction systems, (3) literature analysis tools, (4) writing assistants, and (5) graphic design and reference management solutions. Our assessment framework evaluated key functionalities, costs, technical capabilities, and practical limitations through comprehensive documentation analysis, operational testing, and a systematic review of demonstration materials. All assessments reflect tool versions accessed between October 2024 and March 2025, acknowledging the rapidly evolving nature of this ecosystem. AI technologies primarily enhanced efficiency in literature discovery, content synthesis, and manuscript preparation while maintaining methodological rigor. The 43 evaluated tools demonstrated significant capabilities in processing scientific information, with each category offering distinct advantages for specific research tasks. Findings indicate

[1]Department of Digestive Surgery, Faculty of Medicine, Pontificia Universidad Católica de Chile, Santiago 7850000, Chile.
[2]Institut de Recherche Contre les Cancers de l'Appareil Digestif (IRCAD), Research Institute against Digestive Cancer, Strasbourg 67000, France.
[3]Department of General and Minimally Invasive Surgery, University of Buenos Aires, Buenos Aires 1014, Argentina.
[4]Docencia, Asistencia e Investigación en Cirugía Invasiva Mínima (DAICIM) Foundation, Buenos Aires 1014, Argentina.

Correspondence to: Dr. Brandon Valencia Coronel, Department of Digestive Surgery, Faculty of Medicine, Pontificia Universidad Católica de Chile, Santiago 7850000, Chile. E-mail: brandonvalenciacoronel@gmail.com

substantial time reduction in literature searches, document analysis, and manuscript preparation when properly integrated into research workflows. AI-powered tools demonstrate transformative potential for optimizing surgical research processes, providing significant efficiencies from initial literature search to final publication. Successful implementation requires maintaining a critical balance between technological innovation and fundamental scientific principles, with essential human oversight to prevent overreliance on automation that could compromise critical thinking and analytical skills.

INTRODUCTION

Artificial Intelligence (AI) has grown rapidly in the field of medicine and surgery, covering both clinical applications and research areas[1]. Recent developments in generative AI (GenAI), including large language models (LLMs), have demonstrated their potential to process and generate valuable information in key research stages such as literature review, data analysis, and scientific writing[2-9]. Despite several reviews on AI in clinical decision-making, medical education and simulation[10-18], there remains a clear gap regarding the practical implementation of AI tools in each phase of the medical research process. This work focuses on the practical selection and responsible use of AI platforms within research workflows, emphasizing their application across the various stages of the surgical research and publication pipeline.

Over the past years, the rapid development of LLMs has produced a diverse ecosystem of tools that interact with scientific literature, support document analysis, and assist with evidence synthesis and writing[2-9]. For surgeon-researchers, the adoption of this technology is subject to structural constraints, such as limited protected research time, inconsistent formal training in research methodologies, and the necessity of balancing operative and clinical responsibilities with academic output. In this context, practical, step-by-step guidance on how to select and use AI tools responsibly is particularly relevant. Within the research workflow, these tools generally address initial literature identification and document interaction, with some extending to structured synthesis and reference management for clinical and surgical topics. However, they do not replace critical appraisal, eligibility assessment, or risk-of-bias evaluation. Although they offer efficiency gains in specific steps, these benefits are highly context-dependent and require human verification by surgeons experienced in research methodology.

This narrative review aims to identify and appraise AI platforms that positively influence the execution of surgical medical research. It provides a structured narrative appraisal that maps AI-driven research tools across the surgical research workflow and emphasizes practical functionality and use-case fit. This guide will help surgeons select appropriate AI tools at each research stage, potentially accelerating knowledge generation and improving research efficiency, describing their functions, costs, capabilities, advantages, disadvantages, and accessibility.

LITERATURE REVIEW

We implemented a methodological approach specifically adapted for the evaluation of emerging AI technologies for surgical research, a domain characterized by limited documentation in available scientific literature. Given the accelerated update cycle of these platforms, our findings represent a time-bound snapshot (October 2024-March 2025).

Tool identification

We implemented specific criteria for the systematic selection of tools to include in our analysis. The inclusion criteria comprised: (1) tools that incorporate verifiable AI technologies (machine learning, LLMs, natural language processing, GenAI or related technologies); (2) platforms with direct application in at least one phase of the medical or surgical research process; (3) tools with sufficient technical documentation to enable substantive evaluation; (4) platforms accessible for assessment through free versions, trials, or detailed

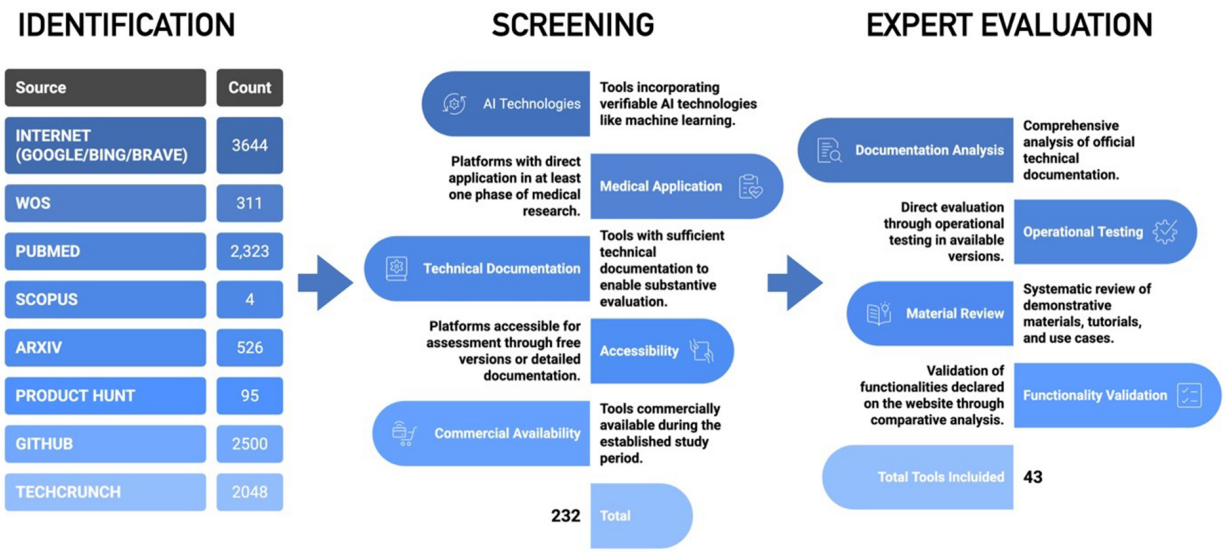


Figure 1. Flowchart outlining the selection process for the tools incorporated within the narrative review.

documentation; and (5) tools commercially available during the established study period. We excluded platforms in beta phase without public availability, systems lacking accessible technical documentation, tools designed exclusively for non-research applications, platforms without distinctive functionalities compared to others already included, and systems discontinued during the evaluation period.

Given the emerging nature of AI-powered tools for surgical research, we adopted a dual search approach. Although exploration in scientific databases (PubMed, Web of Science, Scopus, and arXiv) yielded limited results regarding specific tools, it provided a conceptual framework for the general application of AI technologies in research. We complemented this academic exploration with searches in contemporary technological sources where these emerging innovations are primarily documented, such as general search engines (Google, Bing, Brave) and emerging technology platforms (Product Hunt, GitHub, TechCrunch). This hybrid approach enabled us to capture novel tools that still lack evaluation in the scientific literature but present significant potential for optimizing research processes in medicine.

Evaluation and analysis

Each identified tool underwent a documentation process that included: (1) comprehensive analysis of official technical documentation; (2) direct evaluation through operational testing in available versions; (3) systematic review of demonstrative materials, tutorials, and use cases; and (4) validation of functionalities declared on the website through comparative analysis with similar technologies as seen in Figure 1. This process allowed us to divide the development of AI-enhanced scientific work into 5 categories: academic search engines [Table 1], reading and interaction with documents [Table 2], research and analysis of scientific literature [Table 3], writing assistance systems [Table 4], and graphic design and reference management [Table 5].

In the absence of established frameworks for evaluating AI tools in medical research, we developed a multidimensional evaluative framework that addresses six key domains for practical implementation. This standardized framework examines: (1) Functionalities; (2) Price; (3) Technical capabilities; (4) Operational advantages; (5) Identified disadvantages; and (6) Accessibility. This methodology provides a general assessment specifically adapted to the needs of surgeon-researchers, facilitating informed decisions about the adoption of these emerging technologies.

Table 1. Tools for academic search engines

Model	Functions	Pricing/Month	Capabilities	Advantages	Disadvantages	Accessibility
Consensus AI[19]	Provides evidence-based answers to scientific queries; Summarizes and synthesizes research findings; Extracts key quotes and data from academic papers; Displays confidence scores and citations for transparency	$8.99/Month	Searches across millions of academic studies; Uses natural language processing to generate concise answers; Highlights supporting research and provides direct source links	Quick and intuitive interface for accessing scientific consensus; Saves time by aggregating evidence from multiple studies; Transparent with citations for further verification	Limited to research available in its database; May be less effective for highly theoretical or non-empirical questions; The synthesis accuracy depends on the quality of the underlying papers	www.consensus.app
Elicit[20]	AI-powered research assistant; Automates literature reviews; Summarizes academic papers; Synthesizes research findings	$12/Month	Access to over 125 million academic papers; Provides one-sentence summaries; Extracts data into structured tables; Identifies common themes across papers; Supports PDF uploads for analysis	Saves significant research time; Enhances accuracy in literature reviews; User-friendly interface; Offers both free and paid plans; Integrates with reference management tools	Learning curve for advanced features; Limited to empirical research domains; Accuracy depends on AI interpretations; May require manual verification of data	elicit.org/.com
EvidenceHunt[21]	AI-powered search tool for finding and synthesizing research evidence; Designed to support systematic reviews, policy research, and evidence-based decision making; Retrieves key excerpts and citations from academic and grey literature	€20/Month	Searches across academic publications and online sources; Uses AI to quickly extract and summarize supporting evidence; Provides transparent citations and supporting quotes for verification	Streamlines evidence-gathering for systematic reviews; Saves time compared to manual literature searches; User-friendly interface with a focus on clear, evidence-backed results	Fewer customization options compared to more comprehensive research platforms; The free plan has limited features; Accuracy depends on AI interpretations	www.evidencehunt.com

Manus AI[22]	Literature discovery with semantic expansion and precise query building. De-duplication, screening lists, and inclusion or exclusion tracking. Evidence synthesis with transparent source notes and audit trails. Alerts and periodic refresh to keep surgical evidence up to date. searchable database of over 100 million research papers, and offers highly relevant results	$19/Month	Handles long context for complex surgical topics and heterogeneous outcomes; Generates protocol scaffolds aligned with major reporting standards; Produces publication-ready tables and figures with reproducible templates; Records decision logs for screening and data handling; Integrates discovery, synthesis, analysis, and drafting in one flow	Shortens the time from search to usable evidence for surgical questions; Improves reproducibility through consistent templates and logs; Eases collaboration with role-based sharing and clear handoffs; Speeds manuscript preparation for surgical journals	Requires expert clinical and statistical verification of outputs; Sensitive data must be segregated with formal processing agreements; Risk of incorrect or incomplete citations without librarian review; A credit model can be hard to forecast during heavy screening phases	www.manus.im
OpenRead[23]	AI-powered search across academic papers and web sources; Interactive AI chat assistant ("Oat") for fact-based Q&A; Trending topics display; Integrated reading and note-taking features (e.g., Paper Espresso, Paper Q&A, Related Paper Graph)	$20/Month	Searches a vast repository of over 300 million papers from 20,000+ journals; Allows users to disable web browsing during intensive reasoning tasks; Remembers entire conversation history in a hierarchical tree for easy retrieval	Combines academic and web search to offer deep, up-to-date insights; Enhances research productivity with AI-powered summarization and interactive Q&A; Offers an integrated, all-in-one research environment, including note-taking and trend tracking	Pricing plans are not explicitly detailed on the homepage; The range of advanced features may require a learning curve for new users; The free plan has limited features	www.openread.academy
Perplexity AI[24]	Conversational search engine and Q&A tool; Provides concise, evidence-based answers with online citations; Supports follow-up questions in a natural dialogue; It has a deep research function	$20/Month	Leverages advanced language models for natural language understanding; Retrieves and summarizes information from diverse web sources; Displays transparent source references and citations; Offers a clean, interactive, chat-like interface	Fast, conversational interaction; Transparent with reliable source links; User-friendly with minimal query formatting required; Deep Research mode allows us to perform dozens of searches and read numerous sources to generate a detailed answer	Dependence on the quality of available sources; Limitations in lengthy conversations; Limited premium trial mode; It is not exclusively an academic search engine	www.perplexity.ai

Scinapse[25]	Search and discover academic papers; Filter results by keywords, authors, venues, and more; View citation metrics and related papers; Export citations for further use	Free and $39/Month	Indexes millions of scholarly articles; Provides detailed bibliometric data (e.g., citation counts); User-friendly interface with advanced filtering options	Clean and intuitive design; Fast and free access to academic literature; Useful for quickly finding relevant research	May lack some advanced features found in larger databases (e.g., Google Scholar or Semantic Scholar); Coverage might be less comprehensive in certain niche fields; It doesn't provide AI-generated summaries or conversational responses	www.scinapse.io
Scispace[26]	Search and discover academic papers; It has a deep research function; Automates literature reviews; Supports 75+ languages	$20/Month	Natural-language Q&A inside PDFs with sectioned summaries; Semantic search across 200 M+ papers; Chrome extension & API for external workflows; Semantic filters (topic, study type, date, method)	Saves significant research time; Free tier available for individual researchers and students; Every answer is linked back to the sentence in the original paper, so you can check the source immediately	Certain advanced functionalities (e.g., extended semantic filters and large-scale export) are confined to the Pro tier; Full-text answers depend on whether the journal allows open access; The AI's summaries are helpful but still need human fact-checking	www.scispace.com
Scite AI[27]	Advanced AI-powered citation analysis; Generates Smart Citations (supporting, contrasting, or mentioning); Pro users get unlimited assistant chats and literature searches	$12/Month	Access to over 1.2 billion citations from more than 200 million academic sources; Classifies citations as supporting, contrasting, or mentioning; Integration with reference managers such as Zotero and Mendeley	Provides detailed context on how articles are cited; Offers real-time alerts on new citations and retractions; User-friendly interface; 7-day free trial available	Complex user interface with a steep learning curve; Full access requires a paid subscription; Effectiveness limited by citation data comprehensiveness; Fewer integrations with additional research tools	scite.ai
YouAI[28]	YouChat 2.0 mode answers questions in natural language and displays links to the sources of each piece of information; ARI (Advanced Research & Insights) agent scans; Multilingual chat (> 20 languages) and multimodal prompts keep the same interface for literature	$20/Month	Domain controls mitigate predatory-source noise, improving scholarly precision; Multimodal chat reduces context-switching (code, figures, text in one window)	Citation-linked answers accelerate verification; Research Mode summarizes PDFs/Drive files into section-linked digests, accelerating literature digestion; Custom Model Selector lets researchers compare GPT-4, Claude, Gemini, *etc.*	Full access requires a paid subscription; Learning curve for advanced features; ARI is unavailable to individual researchers unless contracted via enterprise license. AI summaries, while cited, still demand human verification for methodological rigor	you.com

System Pro[29]	Helps researchers find, synthesize, and contextualize scientific literature; Quickly locates useful studies with concise summaries and key details; Exports findings as structured data Generates transparent syntheses from peer-reviewed studies	Unlimited institutional access, price not public reference	Searches through 35 million studies (via PubMed); Offers detailed citation data (e.g., average synthesis cites 53 studies); Refreshes research data daily; Clusters findings into meaningful groups and (coming soon) visualizes topic maps	Reliable and accurate synthesis based exclusively on peer-reviewed research; Significant time-saver for literature review; Transparent with detailed source citations; Provides systemic insights by relating topics within a broader research context	Currently focused on health and life sciences; Some synthesized sentences may require further clarification; Limited to information available in published academic studies	www.system.com/product/system-pro

AI: Artificial intelligence; API: application programming interface.

All the authors with expertise in AI and medical research independently tested each tool. Any discrepancies (e.g., regarding usability or perceived accuracy) were addressed through consensus meetings until agreement was reached, aiming to reduce individual bias. No formal inter-rater statistics were calculated given the narrative design; future studies should include blinded multi-reviewer validation with agreement metrics.

Methodological limitations

We acknowledge several limitations in our approach. First, the inherent dynamism of the technology sector means that our findings represent only a temporal snapshot of the ecosystem. Second, the depth of evaluation varied according to the accessibility of each tool, being more comprehensive for those with complete trial versions. Third, the limited academic literature on these emerging technologies restricts triangulation with previous scientific evidence. Finally, our functional classification is based on subjective analysis, as information was extracted from the platforms themselves and from researcher assessments.

Despite these limitations, this methodological approach constitutes a comprehensive and methodical evaluation of AI-powered tools for optimizing surgical research processes, offering a structured foundation for practical implementations and subsequent investigations.

AVAILABLE AI PLATFORMS

Following a comprehensive analysis of AI tools available for optimizing the surgical research process, a total of 43 platforms have been selected and categorized according to their primary functionality in the scientific workflow. Below, the findings are presented and organized by functional categories.

Academic search engines

Academic search engines with AI integration represent a significant advancement compared to traditional search engines. These tools not only retrieve relevant documents but also synthesize findings from multiple studies, providing evidence-based answers and extracting key citations. As shown in Table 1, platforms such as Consensus AI, Elicit, and Evidence-Hunt stand out for their ability to process natural language and generate concise answers with confidence scores and links to original sources.

Enabling reading and dialogues with documents

Document reading and interaction platforms [Table 2] transform the information processing of complex scientific texts. Tools such as AskYourPDF, ChatPDF, and HumataAI enable conversational dialogues with

Table 2. Tools for enabling reading and dialogues with documents

Model	Functions	Pricing	Capabilities	Advantages	Disadvantages	Accessibility
AskYourPDF[30]	AI-powered assistant for interacting with PDF documents; Allows users to chat with, summarize, and extract insights from PDFs; Offers additional features such as a Chrome extension, Zotero plugin, and mobile apps for iOS and Android; References answers to the original PDF and supports non-PDF formats (Word, PPT, Markdown)	$14.99/Month	Upload and process PDF documents in real time; Generate summaries, answer questions, and extract key quotes; Manage and organize your document library; Multi-file chat and multi-language support	Streamlines the process of extracting information from PDFs; User-friendly and interactive interface; Saves significant time for academic and professional work; Multi-platform availability enhances accessibility	Advanced features may require a premium subscription; Accuracy and depth depend on the quality of the uploaded PDF and underlying AI; Privacy of sensitive documents should be considered	askyourpdf.com
ChatPDF[31]	Enables users to interact with PDF documents conversationally; Allows uploading of any PDF and asking natural language questions about its content; Provides instant answers and summaries based on the document's text	$8/Month	Uses advanced AI (language models) to parse and understand PDF content; Supports dialogue-style Q&A about the document; Handles various PDF formats and sizes	User-friendly and intuitive interface; Speeds up information retrieval from complex documents; No technical expertise required to extract insights	May struggle with very large or highly complex PDFs; Reliant on the quality of the PDF text extraction; Privacy considerations for sensitive documents	chatpdf.com
Humata AI[32]	Conversational interface for interacting with PDF and text documents; Enables users to ask questions, get summaries, and extract insights from uploaded files; Provides context-aware answers and supports document review	$9.99/Month	Upload and process PDF files in real time; Uses advanced language models (e.g., GPT-4) for Q&A and summarization; Maintains conversational history for follow-up queries; Supports multiple document interactions	User-friendly, chat-based interface simplifies information retrieval; Saves time by quickly extracting key insights and summarizing complex documents; Helps users navigate lengthy or technical documents with ease	Free tier may have limitations on document size or usage; Occasional inaccuracies or incomplete responses may occur; Advanced features require a paid subscription	humata.ai

Myreader AI[33]	AI-powered reading assistant that lets you upload and interact with various document formats (PDFs, EPUBs, Kindle, Word, PowerPoint, web articles, and YouTube videos); Provides an interactive chat for Q&A and summarization across your library; Offers smart citations to jump to exact pages in your documents; Organizes your documents into collections accessible from any device	$15/Month	Supports universal document uploads and cloud storage; AI chat for generating summaries, answering questions, and extracting insights; Text-to-speech conversion and multi-language support; Tools to manage and organize your digital library with privacy and security guarantees	Consolidates multiple reading and research features in one platform; Saves time by providing quick summaries and direct access to cited sections; Enhances study, research, and professional reading experiences	Kindle integration may require additional steps; Complex user interface with a steep learning curve; Upload/page limits might restrict very large libraries; The pricing shows a 20,000-page limit for the free version	myreader.ai
NotebookLM[34]	Helps users compile, organize, and synthesize research from web sources into interactive notebooks; Allows conversational Q&A with your research materials and generates summaries; It enables users to interact with Multiple PDF documents	Free	Integrates with Google search to import and summarize web pages; Uses advanced generative AI for context-aware note-taking and conversational insights; Organizes research content into editable, structured notebooks that save conversation history	Leverages Google's robust search and AI technology; Streamlines research and note-taking processes; Free and seamlessly integrated with your Google account; Interactive, user-friendly conversational interface; Podcast creation of an upload document	Still experimental; may have limited availability or features; Summarization accuracy and advanced capabilities may improve over time; Currently optimized for web-based content and may have limited support for other document types	www.notebooklm.google
Scholarcy[35]	AI-powered summarization and reading assistant for research articles; Extracts key points, highlights figures, tables, and citations; Generates flashcards and summary reports to aid comprehension	$7.49/Month	Uses advanced NLP to condense long academic papers into digestible summaries; Identifies important sections, references, and metadata; Integrates with reference managers such as Zotero	Saves time by quickly summarizing lengthy documents; Enhances reading comprehension for busy researchers; Provides a clear, structured overview of key insights	Summaries may sometimes overlook important details in complex texts; Advanced features require a paid subscription	scholarcy.com

Anara[36]	AI-powered tool for clarifying and simplifying complex text; Provides explanations, summaries, and interpretations of academic or technical content; Designed to "unriddle" difficult language for easier comprehension	$12/Month	Uses advanced language models to break down complex concepts; Generates clear, concise explanations and summaries; Tailors output based on user queries and context	Enhances understanding of complex documents and academic texts; User-friendly interface that simplifies technical language; Flexible output suitable for academic, professional, or personal use	May require manual verification for accuracy in highly specialized fields; Free tier may have usage limitations compared to premium plans	unriddle.ai

AI: Artificial intelligence.

PDF and other types of documents, facilitating efficient extraction through natural language questions. These systems implement advanced language models to analyze content in real time, generating summaries and answering specific questions, significantly reducing document review time.

Research and analysis of scientific literature

Scientific literature analysis tools [Table 3] represent a methodological advancement in bibliographic reviews. Platforms such as Connected Papers, Inciteful, and Litmaps are characterized by their ability to visualize citation networks and conceptual connections between publications, facilitating the discovery of non-evident relationships. These tools employ algorithms that map similarities based on both citation patterns and textual content analysis, providing visual representations that enable researchers to quickly comprehend the scientific landscape in specific fields.

Writing assistance systems

Writing support systems and linguistic models [Table 4] represent a significant advance in the scientific writing process. Platforms such as Aithor, CaktusAI, JenniAI, and models such as ChatGPT, ClaudeAI, and Google Gemini provide contextual support for content creation and refinement. These tools implement advanced models that make suggestions based on the scientific context, facilitate text analysis, support the writing and editing of research papers, and provide suggestions on vocabulary, structure, and grammar. Highlights the integration of grammar checking, plagiarism detection, and citation generation.

Graphic design and reference management

The final group analyzed [Table 5] includes graphic design and bibliographic management tools. Applications such as Map This, Napkin AI, and Visily AI transform text into visual representations, facilitating the communication of complex information through automatically generated diagrams and concept maps. While traditional bibliographic management solutions such as EndNote, Paperpile, and Zotero continue to dominate, there is a growing integration of AI capabilities to improve data extraction and citation in platforms such as Scispace.

Ethical use of AI tools

The use of GenAI in manuscript preparation requires adherence to principles of transparency and accountability established by various editorial boards. As a general rule, authors should explicitly disclose the AI tool used, including its version and access date, clearly specify the tasks performed by the AI (e.g., language editing only), and critically review all outputs prior to incorporating them into the manuscript, as emphasized by COPE (Committee on Publication Ethics) and WAME (World Association of Medical Editors)[63,64]. The updated ICMJE (International Committee of Medical Journal Editors) recommendations

Table 3. Tools for research and analysis of scientific literature

Model	Functions	Pricing	Capabilities	Advantages	Disadvantages	Accessibility
Connected Papers[37]	Generates a visual "graph" of academic papers related to a seed paper; Helps discover and explore citation-based relationships and research networks; Provides detailed metadata for each paper with links to sources	$10/Month	Uses advanced algorithms to map paper similarity and citation connections; Displays an interactive, easy-to-navigate graph; Offers essential paper details and links for further exploration	Intuitive visualization that simplifies literature discovery; Great for quickly understanding the research landscape around a paper	Limited filtering and customization options compared to more comprehensive systematic review tools; Focused mainly on citation networks; may not capture all nuances of research topics	connectedpapers.com
Inciteful[38]	Visual literature discovery tool; Maps citation networks and related research papers; Helps identify influential works and uncover hidden connections in scientific literature	Free	Generates interactive graphs from a seed paper; Uses advanced algorithms combining citation and text similarity analyses; Enables exploration of both forward and backward citation relationships; Supports iterative literature exploration and discovery	Intuitive visual interface for quickly grasping the research landscape; Simplifies discovery of relevant papers; Completely free to use; Useful for both initial literature surveys and in-depth analysis	Requires JavaScript enabled for full functionality; May have limitations in filtering or customization compared to more comprehensive systematic review tools; As a relatively new tool, some advanced features may still be under development	inciteful.xyz
Litmaps[39]	Visual literature mapping and discovery; Explore citation networks and related research articles; Track emerging trends and stay updated on new publications; Create and share personalized literature maps	$10/Month	Interactive, graph-based visualization of academic papers and their connections; Custom alerts for new related works; Easy organization and export of literature maps Supports iterative exploration of research fields	Intuitive visual interface that enhances literature discovery; Facilitates identification of influential papers and emerging trends; Useful tool for systematic reviews and comprehensive literature mapping; Easy sharing and collaboration on literature maps	May require a learning curve for users new to graph-based interfaces; Quality of mapping depends on the underlying citation data; Some advanced features are available only in premium plans	www.litmaps.com

Research Rabbit[40]	Discover and explore scientific literature; Visualize networks of research papers, authors, and citations; Track research topics and identify emerging trends; Organize literature collections and create interactive graphs	Free	Graph-based literature discovery and exploration; Seamless navigation through connected research; Real-time updates as new papers and connections emerge; User-friendly interface designed for iterative exploration	Enhances literature review with intuitive visual maps; Facilitates discovery of unexpected connections among research articles; Completely free and accessible; Supports dynamic, exploratory research workflows	May not cover every discipline or database exhaustively; Advanced filtering options can be limited compared to some traditional bibliographic databases; As with any visual tool, performance depends on the quality of underlying data sources	www.researchrabbitapp.com

require this information to be included in both the cover letter and the manuscript itself and explicitly prohibit chatbot authorship[65]. The JAMA Network also recommends verifying the accuracy of citations and maintaining a record of relevant prompts for ethical audit purposes[66]. Major publishing groups, such as *Elsevier* and *Springer Nature*, require protecting sensitive data and suggest uploading conversation logs as supplementary material when appropriate[67,68]. In alignment with these guidelines, we propose a structured set of best practices for the responsible use of AI-based tools in scientific research in Figure 2.

THE CHANGING LANDSCAPE OF RESEARCH IN THE ERA OF AI

The advent of AI has generated a paradigmatic transformation in both medical and surgical research, optimizing workflows that have historically been intensive in human resources and time. Recent studies demonstrate that AI technologies, primarily LLMs, machine learning algorithms, and natural language processing are being applied with promising results across various stages of the research process[2-4,6,7,69,70]. This technological integration addresses the need to manage growing volumes of scientific literature that exceed human processing capacity, allowing researchers to accelerate their analytical processes without compromising methodological rigor.

The findings of our review demonstrate that the incorporation of AI-powered tools could enhance surgical research, from initial literature searches to the presentation and communication of results. Nevertheless, it is essential to analyze both the opportunities and risks associated with each category of tools, considering their impact on the quality, efficiency, and rigor of scientific production. Fabiano *et al.* (2024)[9] highlight in a methodological review that numerous AI applications can streamline each stage of the systematic review process. These tools assist in developing search strategies, automating title/abstract screening, extracting key data, and summarizing findings, thereby markedly improving efficiency.

Academic search engines such as Consensus, Elicit, and EvidenceHunt [Table 1] constitute an initial interface to the scientific corpus. They accelerate scientific searches and increasingly provide LLM-assisted, source-linked synthesis that contextualizes results. However, these tools do not perform study screening, nor do they replace critical appraisal and eligibility assessment. To illustrate how GenAI workflows can affect the subsequent stages of the research pipeline, Trad *et al.*[71] evaluated an LLM-based workflow combining prompt engineering and retrieval-augmented generation. They applied this workflow to a completed systematic review on vitamin D and falls, which included 14,439 records. This workflow reduced screening time from 564.4 to 25.5 hours (a 95.5% reduction), achieved a 0% false-negative rate, and demonstrated 99.6% specificity. Thus, it outperformed manual screening in this context.

Table 4. Tools for writing assistance systems

Model	Functions	Pricing	Capabilities	Advantages	Disadvantages	Accessibility
Aithor[41]	AI-powered research and writing assistant; Helps generate and refine academic or professional content; Supports research tasks including literature synthesis and drafting assistance	$12.99/Month	Utilizes advanced language models for context-aware writing assistance; Assists with research by synthesizing information and generating drafts; Designed for both academic and professional writing needs; Offers integration with research workflows via its web interface	Enhances productivity by streamlining research and writing tasks; Provides contextual, AI-driven content suggestions; Suitable for diverse research disciplines and professional content creation	Free plan has limited features; The AI's outputs are helpful but still need human fact-checking	aithor.com
CaktusAI[42]	AI-powered writing and research assistant; Helps draft, edit, and summarize academic or professional content; Optimize the academic writing process; Generates structured essays in APA/MLA/Chicago with inline citations	$19.99/Month	Uses advanced language models for context-aware text generation and summarization; Supports idea generation and content refinement; Speech-to-text drills and translation for 10 + languages	Streamlines the writing process and boosts productivity; User-friendly interface tailored for academic/professional use; Single hub for writing, coding and math lowers tool-switching friction; Produces full essay skeletons that save hours on structure	May require human review for detailed editing; Some advanced features are locked behind a subscription plan; Academic-integrity concerns (AI-detector evasion)	caktus.ai
ChatGPT[43]	Assists in drafting and editing research papers; Generates summaries and introductory sections; Offers suggestions for vocabulary, structure, and grammar; Answers complex questions and clarifies concepts; Conducts multi-step research through web browsing	$20/Month	Processes and generates coherent, contextually relevant text; Summarizes scientific articles, providing overviews of findings and conclusions; Utilizes the "deep research" tool to autonomously browse the web and compile comprehensive reports; Interprets and analyzes text, images, and PDFs; Plans to include visualization capabilities in future updates; Multimodal processing of text, code, images, audio, and video	Saves time in drafting and reviewing scientific documents; Enhances text quality and coherence; Streamlines complex research tasks, reducing hours of work in minutes; User-friendly interface accessible to both experts and novices	May produce inaccurate information or "hallucinations" requiring verification; Does not replace the critical judgment and expertise of researchers; Ethical debates exist regarding its use in scientific authorship; Academic use is not the only application for this tool	chatgpt.com

ClaudeAI[44]	Assists in drafting and editing research papers; Summarizes complex documents; Answers to complex research questions; Analyzes data and code; Provides real-time image analysis	$20/Month	Advanced natural language processing; Code generation; Integration with various applications via API; Multimodal processing of text, code, images, audio, and video; "Extended Thinking" with web-search, code-exec & MCP connectors for tool-augmented chain-of-thought	Enhances efficiency in research tasks; User-friendly interface; Mobile accessibility through dedicated apps; Native Google Workspace integration streamlines citation & reference insertion	Requires human oversight for critical applications; May produce inaccurate information or "hallucinations" requiring verification; Pro tier enforces 45-message/5-hour quota, interrupting extended writing sessions	claude.ai
Cohere[45]	Drafts, expands and paraphrases sections of manuscripts, policy briefs and technical reports; Includes ready-made connectors for Google Drive, Slack and 80 + workplace apps; Uses Cohere's Command models to write natural-language replies	Free	Powered by Cohere's Command LLM; Customizable with proprietary company data; Ensures data privacy and security	Requires integration with existing enterprise systems; Flexible API lets teams plug the model into custom LaTeX or reference-manager workflows; Continuous updates may necessitate ongoing maintenance	No image-vision modality (text-only); May produce inaccurate information or "hallucinations" requiring verification; This is most beneficial in organizational research environments rather than for individual researchers; Costs rise with very long writing sessions cohere	cohere.com
DeepSeek[46]	Assists in research by generating coherent and relevant text; Answers complex questions and clarifies concepts; Solves logical and mathematical problems; Generates code and supports programming tasks; The reasoning function is available at no cost	Free	Trained on a dataset of 2 trillion tokens; Available across multiple platforms, including web, mobile apps, and API; Multilingual (Chinese, English ± 20 languages) with competitive maths & code benchmarks	Open-source model allowing customization; Cost-effective development; Rapid adoption across various sectors; Very low cost per token among > 100 B-param LLMs	May require human oversight for critical tasks; Potential data privacy concerns due to the open-source nature; Ethical and geopolitical considerations in deployment; Hallucinates citations, occasionally inserting non-existent references into drafts	chat.deepseek.com

Google Gemini[47]	Assists in drafting and editing research papers; Analyzes complex datasets; "Scheduled actions" deliver daily prompts or weekly summaries inside Gmail/Docs; Gives real-time grammar and style tips in Docs	$20/Month	Multimodal processing of text, code, images, audio, and video; Integration with Google AI Studio and Gemini API; Supports real-time streaming and tool use; Works in 40 + languages, including Spanish	Simple interface that is easy for beginners; Facilitates collaboration across various media formats; Seamless integration with Google's ecosystem; Large context window keeps the whole manuscript visible while editing	Requires human oversight for critical applications; May produce inaccurate information or "hallucinations" requiring verification; Most advanced features require a paid subscription	gemini.google.com
GrokAI[48]	Assists in drafting and editing research papers; Offers "Deep Search" and "Think Mode" for richer, step-by-step answers in the writing process; Pulls real-time posts from X to add up-to-date examples	$40/Month	Advanced natural language processing; "Fun/Regular" modes switch tone; Multilingual support; Deployment across multiple platforms including web, iOS, and Android	Open beta API allows custom integrations; Web chat requires no extra software; runs in any X browser session; Long context reduces fragmentation when working on multi-section manuscripts	Full functions only in premium; May produce inaccurate information or "hallucinations" requiring verification; Available mainly inside X; no native Google Docs/Word plug-in yet; Not suited for technical or academic writing	www.grok.com
JeniAI[49]	Provides real-time AI Autocomplete to extend a sentence or start the next one while you type; Generates literature-review drafts, outlines, summaries and paraphrases; Inserts dynamic in-text citations and auto-builds reference lists in 1,700 + styles	$30/Month	Stores, tags and searches personal PDF library; unlimited uploads on Unlimited tier; Offers plagiarism detection and AI-detection mitigation for academic texts; Supplies mini-tools (thesis-statement generator, summarizer, essay expander, paraphraser); Supports 30 + languages for multilingual manuscript drafting	Purpose-built for academic writing; widely adopted by > 5 million researchers; Citation engine reduces reference errors and accelerates compliance with journal styles; Built-in research library and BibTeX import ease migration from Zotero/Mendeley	Free tier imposes 200-autocomplete and 10-chat limits, unsuitable for full-length manuscripts; Autocomplete may insert generic prose requiring substantial subject-expert editing; Current plagiarism checker flags false positives and lacks cross-publisher coverage	jenni.ai
MetaAI[50]	Summarizes uploaded PDFs and highlights research gaps; Assists in drafting and editing research papers; Rewrites text for clarity and academic tone in 30 + languages; Free web chat at meta.ai answers questions and drafts paragraphs	Free	Models such as LLaMA 3 with up to 405 billion parameters; Advanced natural language processing; Simultaneous translation in 101 languages; Open-source tools and libraries for developers and researchers	Promotion of collaboration within the research community; Integration with Meta platforms such as Facebook Instagram and Whatsapp	Requires human oversight for critical applications; May produce inaccurate information or "hallucinations" requiring verification; No built-in citation manager; references must be added manually	meta.ai

Mistral AI[51]	Le Chat drafts, rewrites and summarizes text in a web chat; Pro adds document upload and code interpreter tools; Enhances multilingual communication; Assists in drafting and editing research papers	Free	Advanced natural language processing; Multilingual generation (English, French, German, Spanish, Italian); 128 k-token context in premier models supports full-length systematic-review drafts	Open-source models promote transparency and collaboration; Customizable solutions tailored to specific research needs; Strong focus on data privacy and security; Active engagement in AI policy and ethical considerations	Potential need for technical expertise to implement and customize models; Pay-as-you-go costs climb unpredictably during long drafting sessions	mistral.ai
Paperpal[52]	AI-powered academic writing and editing assistant; Offers grammar checking, paraphrasing, plagiarism detection, and contextual AI writing support; Provides tools for citation generation, research & reference finding, and submission-readiness checks; Integrates with MS Word, Overleaf, and has Chrome extensions and mobile apps	$11.5/Month	Real-time editing and language checks; Advanced generative AI "Copilot" for drafting and refining academic texts; Comprehensive suite including grammar, paraphraser, plagiarism checker, translator, and citation generator; Supports multiple platforms for a seamless academic workflow	Tailored for academic and professional writing; Consolidates multiple writing and editing tools into one platform; Enhances productivity and manuscript quality	Most advanced features require a paid subscription; AI-generated suggestions may still need human review; Users must consider privacy for sensitive documents	paperpal.com
QuillBot[53]	AI-powered writing assistant for paraphrasing and rewording text; Summarization tool; Grammar and plagiarism checking; Citation generation	$4.17/Month	Offers multiple paraphrasing modes (Standard, Fluency, Creative, *etc.*); Summarizes lengthy documents and articles; Integrates grammar and plagiarism checks; Supports citation management	Enhances text clarity and originality; User-friendly interface with a variety of writing modes; Versatile tool for students, researchers, and content creators	Free version is limited in word count and features; AI-generated suggestions may require manual editing for precision and nuance	quillbot.com
Sudowrite[54]	AI-powered creative writing assistant; Helps generate ideas, dialogue, descriptive prose, and narrative continuations; Offers rewriting and stylistic enhancement tools to refine creative content	$10/Month	Utilizes advanced language models to produce creative prompts and text continuations; Supports brainstorming, rewriting, and expanding existing text; Provides a user-friendly interface designed for creative workflows	Helps overcome writer's block and sparks fresh ideas; Speeds up the drafting process and enhances narrative style; Intuitive design tailored for creative writing	This tool is not designed specifically for academic writing; AI-generated content may require human editing for coherence and context; Output quality can vary depending on input and context	sudowrite.com

QwenChat[55]	Assists in drafting and editing research papers; Generates summaries and literature reviews; Summarizes ultra-long PDFs and auto-generates structured outlines (intro, methods, results); Supports multilingual data analysis	Free	Processes up to 128,000 tokens in context; Supports over 29 languages; Offers open-source flexibility for customization	Enhances efficiency in research tasks; Facilitates collaboration across linguistic barriers; Integrates seamlessly with Alibaba Cloud services	Potential data privacy concerns due to open-source nature; No native Google Docs/Word plug-in; integration work needed; Requires human oversight for critical applications	chat.qwen.ai
Yomu AI[56]	AI-powered writing and research assistant; Helps users understand, summarize, and generate text from documents; Aims to boost productivity by simplifying content comprehension and generation	$19/Month	Processes and summarizes uploaded documents; Provides contextual insights and writing suggestions; Uses advanced language models to assist with research and content creation; Designed for academic, professional, and creative writing tasks	Enhances efficiency by quickly summarizing and generating content; User-friendly interface with clear focus on document comprehension; Versatile tool for multiple writing and research scenarios	Most advanced features require a paid subscription; As with all AI tools, output may require human review to ensure accuracy and context	yomu.ai

AI: Artificial intelligence; APA: American Psychological Association; MLA: Modern Language Association; API: application programming interface; MCP: model context protocol; LLM: large language model.

However, significant limitations persist in these systems[72,73,74]. We observed that, although ChatGPT-4 generates search strategies comparable to those developed by expert librarians, these present variations in specificity that could compromise the thoroughness of systematic reviews. Similarly, Feng *et al.*[75] reported high sensitivity (0.928) but low precision (0.200) in machine learning algorithms for the detection of relevant studies, indicating the need for human supervision in bibliographic selection processes.

Document dialogue platforms [Table 2] such as AskYourPDF, ChatPDF, and HumataAI have emerged as intermediaries between researchers and scientific literature, enabling conversational interactions with complex texts. Although formal peer-reviewed evaluations remain limited, our practical evaluation suggests that these technologies can facilitate efficient extraction of specific information, summary generation, and contextual data retrieval.

Research and analysis of scientific literature platforms [Table 3], such as Connected Papers, Inciteful, and Litmaps, represent significant innovations for understanding the structure of different scientific fields. These platforms allow researchers to identify citation patterns, discover influential works, and visualize the evolution of specific knowledge domains. However, in contexts where conversational LLMs are used to automate scientific literature search and review tasks, Yip *et al.* (2024)[76] caution about limitations related to the accuracy, consistency, and relevance of the results obtained, especially in biomedical literature review scenarios.

Academic writing assistance platforms [Table 4], such as Aithor, CaktusAI, and JenniAI, represent a significant advance in the partial automation of the scientific writing process. These tools facilitate drafting,

Table 5. Tools for graphic design and reference management

Model	Functions	Pricing	Capabilities	Advantages	Disadvantages	Accessibility
EndNote[57]	Manages references and citations; Collects references from academic databases, web pages, and PDFs; Automatically formats bibliographies in numerous citation styles; Integrates seamlessly with word processors (e.g., Microsoft Word); Supports sharing and collaboration among	$249.95 Licence one-time	Organizes references with tagging, groups, and notes; Offers advanced search and filtering options; Manages PDFs with annotation and linking features; Synchronizes libraries across devices and supports online collaboration	Robust, widely adopted tool with comprehensive support and training; Deep integration with academic databases and word processing software; Extensive customization through various citation styles	Steeper learning curve compared to some free alternatives; Cost may be prohibitive for some users; Proprietary software limits customization compared to open-source solutions	endnote.com
Map This[58]	Converts PDFs into structured mind maps; Facilitates real-time collaboration; Offers customizable templates; Supports seamless sharing and export options	$5.99/Month	AI-driven document analysis; User-friendly interface; Integration with cloud storage services; Real-time collaboration	Enhances learning and information retention; Streamlines complex data visualization; Supports team collaboration; Accessible without credit card for trial	Limited export formats; Occasional synchronization issues; Requires internet connection for full functionality	map-this.com
Napkin AI[59]	Converts text into visuals such as diagrams, charts, and images; Offers fully editable visuals with customization options; Supports export in .png, .pdf, or .svg formats; Facilitates real-time collaboration with commenting and editing features	Free during beta period	AI-powered text-to-visual conversion; Extensive customization options; Versatile export formats; Real-time collaboration capabilities	Enhances communication efficiency; User-friendly interface; Supports various use cases including presentations, blogs, social media, and documentation	Free plan has limited features; Best editing experience on desktop; mobile editing may be limited; Requires human oversight to ensure accuracy of generated visuals	napkin.ai
Paperpile[60]	Manages research libraries and references; Organizes PDFs and other documents; Automatically generates citations and bibliographies; Enables highlighting, annotation, and collaboration; Integrates with Google Docs, Microsoft Word, and mobile devices	$35/Year	Cloud-based reference management with seamless integration with Google Drive and Docs; Automatic citation formatting in multiple styles Tools for PDF organization, annotation, and sharing; Supports cross-device access with iOS and Android apps	User-friendly interface designed specifically for researchers; Streamlines citation and bibliography management; Enhances collaboration with team-sharing features; Provides efficient organization of research materials in one centralized location	Limited offline functionality; Advanced features require a paid subscription; May not cover every specialized need for extensive research libraries; Does not work with generative artificial intelligence	paperpile.com

Zotero[61]	Manages references and research materials; Collects citations from the web, databases, and PDFs; Automatically extracts metadata and generates bibliographies Integrates with word processors (Word, LibreOffice) and browsers	$20/Month	Organizes research items with tagging, collections, and notes; Supports various citation styles and automated formatting; Syncs libraries across multiple devices and allows for collaborative group libraries; Offers browser extensions for easy	Completely free and open source; Highly customizable with strong community support; Seamless integration with academic workflows and multiple platforms; Ideal for collaborative research	Free storage limit may be restrictive for heavy users; Interface may feel less modern compared to some commercial alternatives; Advanced customization can require technical know-how	zotero.org/
Visily AI[62]	Converts text, screenshots, or sketches into editable wireframes and prototypes; Offers AI-powered design assistance; Provides a rich library of UI components and templates; Facilitates real-time collaboration among team members	$11/Month	AI-driven design generation; Extensive template and component library; Real-time collaborative workspace; Integration with tools such as Figma	User-friendly interface suitable for non-designers; Accelerates design workflows; Enhances team collaboration; Offers flexible pricing plans	The free plan has limitations on features and credits; Advanced features require a Pro subscription; May require internet connection for optimal performance	visily.ai/

AI: Artificial intelligence; UI: user interface.

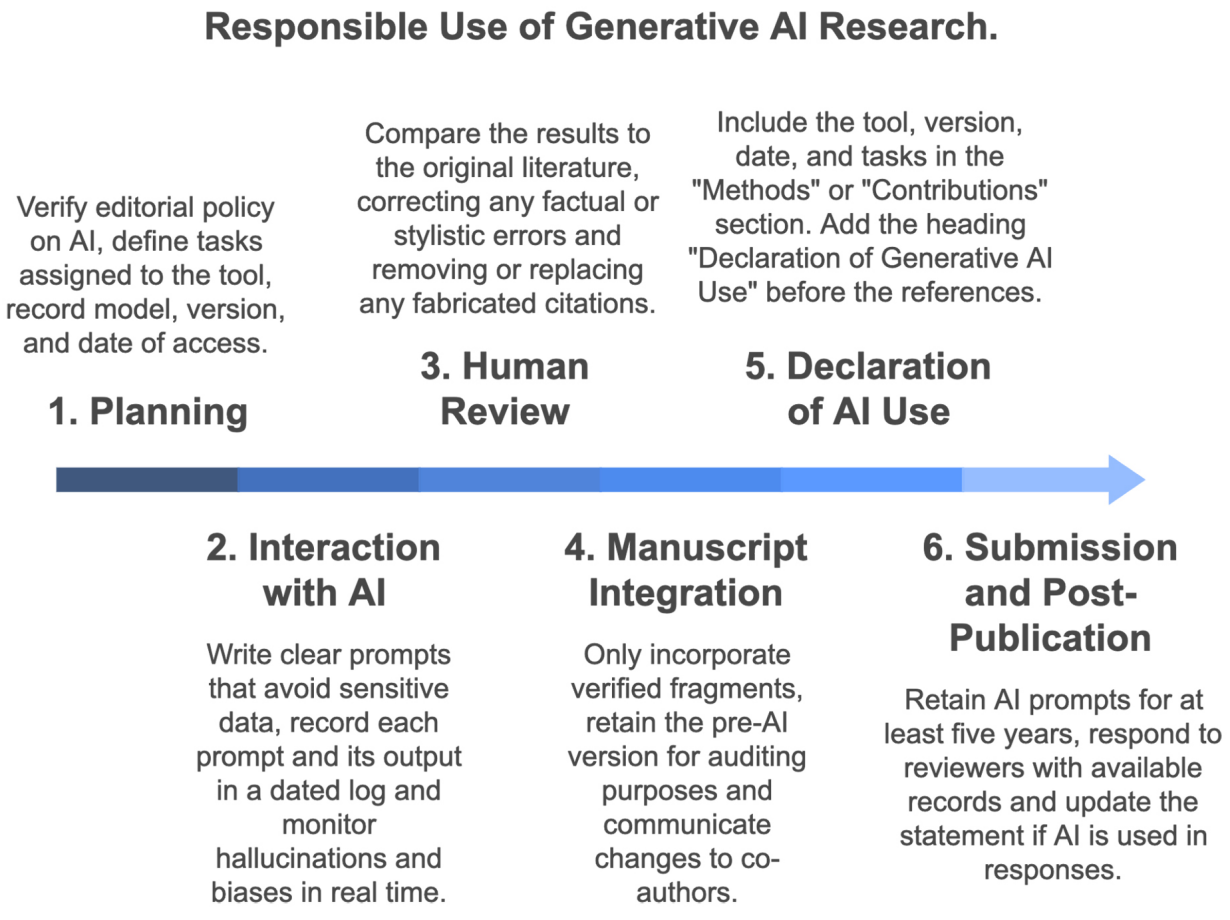


Figure 2. Responsible Use of Generative AI Research. AI: Artificial intelligence.

paraphrasing, and stylistic refinement, thus reducing the time spent on mechanical writing tasks. Esposito *et al.* (2024)[77] demonstrate that tools such as ChatGPT can produce analyses comparable to those performed by human researchers, especially after multiple human interactions with the chatbot, and it also reduces research analysis time.

However, multiple studies emphasize the limitations of these technologies in highly specialized scientific contexts. Tercero-Hidalgo *et al.* (2022)[78] identified that the use of AI in evidence synthesis for coronavirus disease 2019 (COVID-19) lacked methodological consistency. These findings emphasize the importance of human supervision and validation of automatically generated content that answers intricate research queries.

Data visualization and bibliographic management tools [Table 5] facilitate effective communication of scientific findings and systematic organization of references. Platforms such as Map This, Napkin AI, and Visily AI allow for the transformation of complex data into accessible visual representations, facilitating the comprehension of results [57-62]. Although specific evidence on the impact of these tools in surgical research is limited, Raja *et al.* (2023)[79] demonstrated the effectiveness of LLMs for automating the analysis of categories and trends in ophthalmology scientific articles, suggesting potential applications in dynamic data visualization.

Reference managers, although they do not incorporate advanced AI technologies in all cases, constitute essential components of the modern research workflow. Their integration with other AI-powered tools represents a significant opportunity for the holistic optimization of the research process.

Beyond efficiency and rigor, GenAI tools may also act as catalysts for creativity in research. Fisher *et al.* (2025) [80] found in a randomized trial that students using ChatGPT exhibited enhanced creative output (idea fluency and elaboration) in certain fields, although those with already high intrinsic creativity showed little benefit. This implies that AI can serve as a valuable creativity scaffold for users who need inspiration or diverse perspectives, while for more inherently creative individuals, it shifts to a collaborative or productivity-enhancing role. Thoughtfully integrating AI in tasks such as hypothesis generation or study design could thus stimulate innovative thinking, aligning with our focus on how AI might augment the creative aspects of surgical research.

In surgical research, the adoption of AI extends beyond efficiency; it represents a paradigm shift in how surgeons generate, interpret, and disseminate scientific knowledge. AI tools can help surgical teams transform raw information into structured, publishable evidence. As presented in this review, they provide a foundation for integrating AI into the entire surgical research ecosystem. By empowering surgeons to optimize literature synthesis, data management, and manuscript preparation, these technologies have the potential to accelerate discovery, enhance global collaboration, and strengthen the academic impact of surgical innovation.

We must avoid the loss of critical thinking, particularly among early-career researchers who might rely excessively on AI to the point of delegating core scholarly tasks. While AI can be a powerful accelerator, indiscriminate use risks eroding the skills required for rigorous research, analysis, and writing. Emerging evidence cautions against LLM-first writing: in an essay-writing paradigm, Kosmyna *et al.* (2024)[81] reported that participants using continuous LLM assistance exhibited weaker neural engagement and lower content recall than those writing unaided; with repeated LLM-only use, performance declined relative to controls, consistent with an accumulation of "cognitive debt." By contrast, a brain-first, AI-second sequence, independent conceptualization and drafting followed by targeted LLM refinement, was associated with higher neural connectivity (on the order of 50%-200% increases across canonical electroencephalogram (EEG) bands) and preserved recall. Although our review is descriptive and does not assess cognitive

outcomes for individual tools, these preliminary findings support our recommendation for balanced, staged AI use with sustained human oversight and source-linked verification to protect critical thinking while leveraging efficiency gains.

Finally, broad empirical evidence supports the notion that AI tools can improve learning. A recent meta-analysis of 51 studies reported that using ChatGPT yielded a large positive effect on overall performance ($g \approx 0.867$) and moderately improved users' learning perception and higher-order thinking skills[82]. Although focused on educational settings, these results suggest that appropriately integrated AI assistance can enhance not only efficiency but also the quality and depth of cognitive work. This aligns with our argument that AI, when used judiciously, can bolster both the productivity and the rigor of surgical research endeavors.

LIMITATIONS

This review has inherent methodological limitations that should be considered when interpreting the results. A fundamental challenge is the extraordinarily dynamic nature of the ecosystem of AI-based tools for surgical research. The technological landscape is evolving at an accelerated pace, with frequent feature updates, changes in pricing models, and the continuous introduction of new platforms. As a result, this analysis inevitably represents a snapshot in time that can change significantly over short periods of time.

A second major limitation is the potentially subjective nature of the ratings presented. While a structured methodology has been implemented to categorize and analyze the tools, the information is primarily derived from the platforms themselves and the experts' assessments, introducing a degree of subjectivity and potential bias. The lack of established standards for evaluating AI-based technologies in surgical research made it difficult to develop fully objective metrics for performance and utility.

This assessment was primarily qualitative. We did not employ pre-specified quantitative benchmarks (e.g., exact time-savings or citation-accuracy scores), given the lack of standardized evaluation frameworks for AI-powered research tools. Instead, we focused on qualitative indicators - perceived efficiency of literature search, transparency of citations, usability, and consistency of outputs - during hands-on testing, interpreting our findings as expert-based descriptive assessments rather than outcomes of controlled quantitative experiments. No numerical weighting was assigned across the six evaluation domains. All domains were qualitatively considered. In case of trade-offs, we prioritized methodological transparency and verifiability (e.g., citation traceability), functionalities with direct impact on research rigor, and cost/accessibility for resource-limited settings.

In addition, the depth of evaluation varied significantly between tools, depending on the accessibility of trial versions and the availability of technical documentation. This may have led to inconsistent characterizations, particularly for platforms with limited free access. The lack of peer-reviewed scientific literature specifically focused on these emerging technologies also limited the ability to triangulate our findings with previous research.

CONCLUSION

AI-powered tools demonstrate a transformative potential for optimizing surgical research processes, providing significant efficiencies in tasks involving extensive literature review, literature data analysis, and research writing. The successful implementation of these tools, however, requires a critical balance between technological innovation and the fundamental principles of scientific rigor and expert judgment that underpin excellence in surgical research. In the coming years, surgeons must prepare for the increased integration of AI into standard research infrastructure, such as reference managers, selection workflows and journal submission requirements. In the rapidly evolving field of GenAI, the preparation of local policies,

training in AI skills, and the maintenance of human critical evaluation will be essential. We believe in the critical importance of maintaining human oversight to prevent overreliance on automation that could potentially compromise critical scientific thinking and analytical skills. Consequently, future studies should examine both the advantages and potential risks of AI in scientific inquiry because these platforms are frequently updated, feature sets and pricing should be interpreted as time-bounded and may require periodic reassessment.

DECLARATIONS

Authors' contributions

Conceptualized the review, conducted the initial analysis, drafted the manuscript, performed the primary writing, and provided expert analysis and a critical review of the manuscript: Valencia-Coronel B
Conceptualized the review, contributed to the systematic evaluation of AI tools, corrected the manuscript, and provided expert analysis and a critical review of the manuscript: Marescaux JF, Gimenez M
All authors participated in the consensus-building process for tool evaluation, contributed to the interpretation of findings, critically reviewed and revised the manuscript, and approved the final version for publication.

Availability of data and materials

This work did not generate any new data. All data supporting the conclusions of this review are included in the published literature cited in the reference list.

Financial support and sponsorship

None.

Conflicts of interest

Marescaux JF serves as an Honorary Regional Editor of *Artificial Intelligence Surgery*, and Valencia-Coronel B is a member of the Junior Editorial Board of *Artificial Intelligence Surgery*. Neither Marescaux JF nor Valencia-Coronel B was involved in any part of the editorial process for this manuscript, including the selection of reviewers, manuscript handling, or decision-making. Gimenez M declares that he has no conflicts of interest.

Ethical approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Copyright

© The Author(s) 2026.

REFERENCES

1. Meng X, Yan X, Zhang K, et al. The application of large language models in medicine: a scoping review. *iScience*. 2024;27:109713. DOI PubMed PMC
2. Perlis RH, Fihn SD. Evaluating the application of large language models in clinical research contexts. *JAMA Netw Open*. 2023;6:e2335924. DOI PubMed
3. Kothari AN, Kaji AH, Melton GB. Practical guide to the use of AI-enabled analytics in research. *JAMA Surg*. 2025;160:590-1. DOI
4. Loftus TJ, Haider A, Upchurch GR Jr. Practical guide to artificial intelligence, chatbots, and large language models in conducting and reporting research. *JAMA Surg*. 2025;160:588-9. DOI PubMed
5. Jhajj KS, Jindal P, Kaur K. Use of artificial intelligence tools for research by medical students: a narrative review. *Cureus*. 2024;16:e55367. DOI PubMed PMC
6. Guo E, Gupta M, Deng J, Park YJ, Paget M, Naugler C. Automated paper screening for clinical reviews using large language models: data analysis study. *J Med Internet Res*. 2024;26:e48996. DOI PubMed PMC

7. Wang Z, Cao L, Danek B, Jin Q, Lu Z, Sun J. Accelerating clinical evidence synthesis with large language models. *NPJ Digit Med.* 2025;8:509. DOI PubMed PMC
8. Singhal K, Tu T, Gottweis J, et al. Toward expert-level medical question answering with large language models. *Nat Med.* 2025;31:943-50. DOI PubMed PMC
9. Fabiano N, Gupta A, Bhambra N, et al. How to optimize the systematic review process using AI tools. *JCPP Adv.* 2024;4:e12234. DOI PubMed PMC
10. Hashimoto DA, Varas J, Schwartz TA. Practical guide to machine learning and artificial intelligence in surgical education research. *JAMA Surg.* 2024;159:455-6. DOI PubMed
11. Nerella S, Bandyopadhyay S, Zhang J, et al. Transformers and large language models in healthcare: a review. *Artif Intell Med.* 2024;154:102900. DOI PubMed PMC
12. Pressman SM, Borna S, Gomez-Cabello CA, Haider SA, Haider CR, Forte AJ. Clinical and surgical applications of large language models: a systematic review. *J Clin Med.* 2024;13:3041. DOI PubMed PMC
13. Bedi S, Liu Y, Orr-Ewing L, et al. Testing and evaluation of health care applications of large language models: a systematic review. *JAMA.* 2025;333:319-28. DOI PubMed PMC
14. Yang X, Li T, Su Q, et al. Application of large language models in disease diagnosis and treatment. *Chin Med J.* 2025;138:130-42. DOI PubMed PMC
15. Zhang J, Sun K, Jagadeesh A, et al. The potential and pitfalls of using a large language model such as ChatGPT, GPT-4, or LLaMA as a clinical assistant. *J Am Med Inform Assoc.* 2024;31:1884-91. DOI PubMed PMC
16. Belmar F, Gaete MI, Escalona G, et al. Artificial intelligence in laparoscopic simulation: a promising future for large-scale automated evaluations. *Surg Endosc.* 2023;37:4942-6. DOI PubMed PMC
17. Bhuyan SS, Sateesh V, Mukul N, et al. Generative artificial intelligence use in healthcare: opportunities for clinical excellence and administrative efficiency. *J Med Syst.* 2025;49:10. DOI PubMed PMC
18. Tu T, Schaekermann M, Palepu A, et al. Towards conversational diagnostic artificial intelligence. *Nature.* 2025;642:442-50. DOI PubMed PMC
19. Consensus.app. Available from https://consensus.app/ [accessed 14 January 2026].
20. Elicit.com. Available from https://www.elicit.com [accessed 14 January 2026].
21. EvidenceHunt. Available from https://evidencehunt.com/ [accessed 14 January 2026].
22. Manus: the action engine for research. Available from https://manus.im/ [accessed 14 January 2026].
23. OpenRead. Available from https://www.openread.academy/ [accessed 14 January 2026].
24. Perplexity AI. Available from https://www.perplexity.ai/ [accessed 14 January 2026].
25. Scinapse. Finding R&D Trends and Experts Made Simple. Available from https://www.scinapse.io/ [accessed 14 January 2026].
26. Discover, Create, and Publish your research paper. Scispace. Available from https://scispace.com/ [accessed 14 January 2026].
27. AI for Research. scite.ai. Available from https://scite.ai/ [accessed 14 January 2026].
28. You.com. AI for workplace productivity. Available from https://you.com/ [accessed 14 January 2026].
29. System Pro. Available from https://www.system.com/product/system-pro [accessed 14 January 2026].
30. AskYourPDF. AskYourPDF: The best PDF AI chat app. Available from https://askyourpdf.com/es [accessed 14 January 2026].
31. Chatea con cualquier PDF. Available from https://www.chatpdf.com/es [accessed 14 January 2026].
32. Humata. Humata: AI for your files. Available from https://www.humata.ai/ [accessed 14 January 2026].
33. Myreader. Myreader - Let an AI read books, documents, papers for you. Available from https://www.myreader.ai/ [accessed 14 January 2026].
34. Google NotebookLM. Available from https://notebooklm.google/ [accessed 14 January 2026].
35. Scholarcy - Knowledge made simple. Available from https://www.scholarcy.com/ [accessed 14 January 2026].
36. Anara. anara.com. Available from https://anara.com/?from=unriddle [accessed 14 January 2026].
37. Connected Papers. Available from https://www.connectedpapers.com/ [accessed 14 January 2026].
38. Inciteful. Available from https://inciteful.xyz/ [accessed 14 January 2026].
39. Litmaps. Available from https://www.litmaps.com/ [accessed 14 January 2026].
40. ResearchRabbit. Available from https://www.researchrabbitapp.com/ [accessed 14 January 2026].
41. Asistente de Escritura AI. Available from https://aithor.com/es-419 [accessed 14 January 2026].
42. Caktus AI. Available from https://caktus.ai/login?callbackUrl=https%3A%2F%2Fcaktus.ai%2F [accessed 14 January 2026].
43. ChatGPT. Available from https://chat.openai.com/ [accessed 14 January 2026].
44. Claude. Available from https://claude.ai/ [accessed 14 January 2026].
45. Cohere. Available from https://dashboard.cohere.com/ [accessed 14 January 2026].
46. DeepSeek. Available from https://chat.deepseek.com/ [accessed 14 January 2026].

47. Gemini. Available from https://gemini.google.com/ [accessed 14 January 2026].
48. Grok. Available from https://x.ai/grok [accessed 14 January 2026].
49. Jenni AI. Available from https://jenni.ai [accessed 14 January 2026].
50. Meta AI. Meta AI. Available from https://www.meta.ai/ [accessed 14 January 2026].
51. Mistral AI. Available from https://mistral.ai/ [accessed 14 January 2026].
52. Paperpal. AI Academic Writing Tool - Online English Language Check. Available from https://paperpal.com/ [accessed 14 January 2026].
53. QuillBot: your complete writing solution. Available from https://quillbot.com/ [accessed 14 January 2026].
54. Sudowrite - best AI writing partner for fiction. Available from https://sudowrite.com/ [accessed 14 January 2026].
55. Qwen Chat. Available from https://chat.qwen.ai/ [accessed 14 January 2026].
56. Yomu AI. Available from https://www.yomu.ai/ [accessed 14 January 2026].
57. EndNote. EndNote - the best citation & reference management tool. Available from http://www.endnote.com/ [accessed 14 January 2026].
58. Map This. Available from https://www.map-this.com/ [accessed 14 January 2026].
59. Napkin AI. Napkin AI - the visual AI for business storytelling. Available from https://www.napkin.ai/ [accessed 14 January 2026].
60. Paperpile. Available from http://www.paperpile.com/ [accessed 14 January 2026].
61. Zotero. Available from https://www.zotero.org/ [accessed 14 January 2026].
62. Visily. Visily - AI-powered UI design software. Available from https://www.visily.ai/ [accessed 14 January 2026].
63. COPE Council. Authorship and AI tools. Committee on Publication Ethics.2024. DOI
64. World Association of Medical Editors. Chatbots, generative AI, and scholarly manuscripts. Available from https://wame.org/page3.php?id=106 [accessed 14 January 2026].
65. International Committee of Medical Journal Editors. Recommendations. Available from https://www.icmje.org/recommendations/ [accessed 14 January 2026].
66. Flanagin A, Pirracchio R, Khera R, Berkwits M, Hswen Y, Bibbins-Domingo K. Reporting Use of AI in Research and Scholarly Publication-JAMA Network Guidance. *JAMA.* 2024;331:1096-8. DOI PubMed
67. Adapting to AI. *Nat Geosci.* 2024;17:481. DOI
68. Elsevier. Generative AI policies for journals. Available from: https://www.elsevier.com/about/policies-and-standards/generative-ai-policies-for-journals?utm_source=chatgpt.com [accessed 14 January 2026].
69. Adam GP, Davies M, George J, et al. Machine learning tools to (semi-) automate evidence synthesis. Rockville: agency for Healthcare Research and Quality (AHRQ); 2024. Available from https://www.ncbi.nlm.nih.gov/books/NBK614977/ [accessed 14 January 2026].
70. Dennstädt F, Hastings J, Putora PM, Schmerder M, Cihoric N. Implementing large language models in healthcare while balancing control, collaboration, costs and security. *NPJ Digit Med.* 2025;8:143. DOI PubMed PMC
71. Trad F, Yammine R, Charafeddine J, et al. Streamlining systematic reviews: a novel application of Large Language Models. *arXiv* 2024;arXiv:2412.15247. Available from https://arxiv.org/abs/2412.15247 [accessed 14 January 2026].
72. Alaniz L, Vu C, Pfaff MJ. The utility of artificial intelligence for systematic reviews and boolean query formulation and translation. *Plast Reconstr Surg Glob Open.* 2023;11:e5339. DOI PubMed PMC
73. Alber DA, Yang Z, Alyakin A, et al. Medical large language models are vulnerable to data-poisoning attacks. *Nat Med.* 2025;31:618-26. DOI PubMed PMC
74. Tam TYC, Sivarajkumar S, Kapoor S, et al. A framework for human evaluation of large language models in healthcare derived from literature review. *NPJ Digit Med.* 2024;7:258. DOI PubMed PMC
75. Feng Y, Liang S, Zhang Y, et al. Automated medical literature screening using artificial intelligence: a systematic review and meta-analysis. *J Am Med Inform Assoc.* 2022;29:1425-32. DOI PubMed PMC
76. Yip R, Sun YJ, Bassuk AG, Mahajan VB. Artificial intelligence's contribution to biomedical literature search: revolutionizing or complicating? *PLOS Digit Health.* 2025;4:e0000849. DOI PubMed PMC
77. Esposito C, Dell'Omo G, Di Ianni D, Di Procolo P. Human *vs.* ChatGPT. Is it possible obtain comparable results in the analysis of a scientific systematic review? *Recenti Prog Med.* 2024;115:420-5. DOI PubMed
78. Tercero-Hidalgo JR, Khan KS, Bueno-Cavanillas A, et al. Artificial intelligence in COVID-19 evidence syntheses was underutilized, but impactful: a methodological study. *J Clin Epidemiol.* 2022;148:124-34. DOI PubMed PMC
79. Raja H, Munawar A, Mylonas N, et al. Automated category and trend analysis of scientific articles on ophthalmology using large language models: development and usability study. *JMIR Form Res.* 2024;8:e52462. DOI PubMed PMC
80. Fisher R, Gogan T, Williams J, et al. Can ChatGPT enhance business student creativity? *Stud High Educ.* ;2025:1-15. DOI

81. Kosmyna N, Hauptmann E, Yuan YT, et al. Your brain on ChatGPT: accumulation of cognitive debt when using an AI assistant for essay writing task. *arXiv* 2025;arXiv:2506.08872. Available from https://newsletter.lesderniershommes.com/content/files/2025/06/2025-MIT---Brain-debt-ChatGPT_compressed.pdf [accessed 13 October 2025].

82. Wang J, Fan W. The effect of ChatGPT on students' learning performance, learning perception, and higher-order thinking: insights from a meta-analysis. *Humanit Soc Sci Commun.* 2025;12:4787. DOI

Disclaimer/Publisher's Note: All statements, opinions, and data contained in this publication are solely those of the individual author(s) and contributor(s) and do not necessarily reflect those of OAE and/or the editor(s). OAE and/or the editor(s) disclaim any responsibility for harm to persons or property resulting from the use of any ideas, methods, instructions, or products mentioned in the content.

© The Author(s) 2026. Open Access This article is licensed under a Creative Commons Attribution 4.0 International License (https://creativecommons.org/licenses/by/4.0/),which permits unrestricted use,sharing, adaptation,distribution and reproduction in any medium or format,for any purpose,even commercially,as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license,and indicate if changes were made.