Integrative Medicine Research 14 (2025) 101222

Contents lists available at ScienceDirect

Integrative Medicine Research

journal homepage: www.elsevier.com/locate/imr

Education Article

Prompt engineering for generative artificial intelligence chatbots in health research: A practical guide for traditional, complementary, and integrative medicine researchers

Jeremy Y. Ng [a,b,c,d]

[a] Institute of General Practice and Interprofessional Care, University Hospital Tübingen, Tübingen, Germany
[b] Robert Bosch Center for Integrative Medicine and Health, Bosch Health Campus, Stuttgart, Germany
[c] Department of Health Research Methods, Evidence, and Impact, Faculty of Health Sciences, McMaster University, Hamilton, Canada
[d] School of Public Health, Faculty of Health, University of Technology Sydney, Sydney, Australia

ARTICLE INFO

Keywords:
AI chatbots
Prompt engineering
Large language models
Generative artificial intelligence

ABSTRACT

Generative artificial intelligence (GenAI) chatbots powered by large language models (LLMs) are increasingly used in health research to support a range of academic and clinical activities. While increasingly adopted in biomedical research, their application in traditional, complementary, and integrative medicine (TCIM) remains underexplored. TCIM presents unique challenges, including complex interventions, culturally embedded practices, and variable terminology. This article provides a practical, evidence-informed guide to help TCIM researchers engage responsibly with GenAI chatbots through prompt engineering, the design of clear, structured, and purposeful prompts to improve output relevance and accuracy. The guide outlines strategies to tailor GenAI chatbot interactions to the methodological and epistemological diversity of TCIM. It presents use cases across the research process, including research question development, study design, literature searches, selection of reporting guidelines and appraisal tools, quantitative and qualitative analysis, writing and dissemination, and implementation planning. For each stage, the guide offers examples and best practices while emphasizing that AI-generated content should always serve as a starting point, not a final product, and must be reviewed and verified using credible sources. Potential risks such as hallucinated outputs, embedded bias, and ethical challenges are discussed, particularly in culturally sensitive contexts. Transparency in GenAI chatbot use and researcher accountability are emphasized as essential principles. While GenAI chatbots can expand access to research support and foster innovation in TCIM, they cannot substitute for critical thinking, methodological rigour, or domain-specific expertise. Used responsibly, GenAI chatbots can augment human judgment and contribute meaningfully to the evolution of TCIM scholarship.

1. Introduction

Generative artificial intelligence (GenAI) chatbots powered by large language models (LLMs), such as OpenAI's ChatGPT, Google Gemini, Meta AI, and Microsoft Copilot, are increasingly used in health research to support a range of academic and clinical activities.[1,2] These tools offer accessible, cost-effective assistance across disciplines,[1,3,4] including in traditional, complementary, and integrative medicine (TCIM). A key factor in optimizing GenAI chatbot use is prompt engineering, the practice of designing precise and structured input queries. Effective prompts improve output quality, enhance alignment with research standards, and reduce the risk of bias or misinformation.[4-8] TCIM includes a diverse range of therapies including mind–body therapies, biologically based practices, manipulative and body-based practices, energy healing, and whole medical systems.[9] Challenges in this field, including intervention heterogeneity and limited high-quality evidence within certain modalities, can affect artificial intelligence (AI) performance,[10-12] while AI also presents new opportunities to support innovation, interdisciplinarity, and access to underrepresented knowledge systems.[11] This guide is informed by four original studies led by the author (including three large-scale, international cross-sectional surveys of researchers, university students and postdoctoral fellows, and journal editors-in-chief, respectively, and all in the biomedical field), and an audit of GenAI publisher policies, all of which generated rich empirical data surrounding the attitudes, perceptions and policies of GenAI tools across the biomedical research and publishing landscape.[1,13–15] In addition to these studies, recommendations from organizations that provide guidance, standards, and best practices for scholarly publishing, including the International Association of Scientific, Technical and Medical Publishers (STM), the International Committee of Medical Jour-

E-mail addresses: ngjy2@mcmaster.ca, jeremy.ng@med.uni-tuebingen.de

https://doi.org/10.1016/j.imr.2025.101222
Received 13 July 2025; Received in revised form 9 August 2025; Accepted 9 August 2025
Available online 9 August 2025

Table 1
Main types of prompting in GenAI chatbot engineering with examples for traditional, complementary, and integrative medicine (TCIM) research.

Type of Prompting	Definition	Key Feature	Ideal Use Case	Example Prompt Relating to TCIM Research
Zero-shot prompting	The model generates an output without any prior examples provided.	No examples given; relies entirely on model knowledge.	When you want a quick, direct answer or generation without context.	"Summarize the main research challenges in traditional herbal medicine."
One-shot prompting	The model is given exactly one example to guide its response.	Single example provided to illustrate desired output.	When the task is specific and a single example can clarify expectations.	"Here is an example of a research abstract on acupuncture efficacy: [abstract]. Now, summarize this new abstract on acupuncture safety."
Few-shot prompting	The model is provided with a few examples before generating output.	Multiple examples to show patterns or formats.	When complex or nuanced tasks require demonstration for accuracy.	"Here are three examples of complementary and alternative medicine (CAM) clinical trial designs: [examples]. Design a trial protocol for a herbal remedy for anxiety."
Chain-of-thought prompting	The model is guided to reason step-by-step in the response.	Encourages explicit reasoning and intermediate steps.	Useful for multi-step problems or complex reasoning tasks.	"Explain step-by-step how to critically appraise a systematic review on yoga for chronic pain."
Role-based prompting	The model is asked to assume a specific persona or role.	Contextualizes output by simulating a role or perspective.	When you want the model to produce domain-specific or stylistically consistent content.	"You are a TCIM journal editor. Provide advice to authors on improving the reporting quality of herbal medicine studies."

nal Editors (ICMJE), the World Association of Medical Editors (WAME), the Committee on Publication Ethics (COPE), and the European Association of Science Editors (EASE),[16–20] were reviewed. Finally, a search of the academic and grey literature was conducted to identify additional guidance relevant to the integration of geneGenAI tools into the health research process. Drawing collectively from these sources, this guide offers a practical, evidence-informed resource for prompt engineering tailored to TCIM researchers.

While several guides on GenAI and prompt engineering exist within the broader medical and scientific literature, they have been published in journals that may not reach TCIM researchers. This guide is tailored specifically for the TCIM research community, many of whom may be less familiar with GenAI tools and concepts. By providing accessible, field-relevant examples, it aims to bridge this gap and support the responsible and effective integration of GenAI into TCIM research workflows.

2. Background: GenAI chatbots and their use in medical research

GenAI chatbots powered by large language models (LLMs) are increasingly used in health research due to their ability to generate coherent, contextually relevant responses to natural language prompts. Trained on large text corpora and built using transformer-based architectures, LLMs can perform complex language tasks that once required human expertise.[21-24] Unlike traditional software, LLM-based chatbots offer conversational interfaces that allow users to interact using plain language, making them accessible to researchers with minimal technical training.[25] These tools are now being applied across all stages of health research, including study design, literature synthesis, and dissemination.[1] Editors and peer reviewers are also exploring their use for preliminary manuscript assessment.[14] However, AI-generated outputs may include errors, fabricated citations, or misleading information, which requires careful scrutiny.[26-28] As such, researchers must consistently apply critical thinking and remain responsible for the accuracy, rigour, and ethical standards of their work if using such tools.[29,30] As the role of GenAI chatbots grows in both conventional and TCIM research, there is an urgent need to educate researchers about prompt engineering to ensure effective and responsible use.

3. What is prompt engineering?

Prompt engineering refers to the strategic design of input queries that guide GenAI chatbot responses generated by large language models (LLMs).[23,31,32] Effective prompts are clear, purposeful, and contextually appropriate, allowing researchers to obtain accurate and relevant outputs. Even small changes in prompt wording or structure can affect outputs, making thoughtful prompt design critical.[5-8]

Several prompting strategies are commonly used. Zero-shot prompting involves giving no examples and is best suited for simple queries.[23,33] One-shot and few-shot prompting includes examples to improve consistency.[23,32,33] Chain-of-thought prompting encourages stepwise reasoning, useful in analytical tasks.[23,34] Role-based prompting asks the GenAI chatbot to adopt a specific role, such as "systematic review expert," to shape tone and response quality.[35,36] A table summarizing these main types of prompting strategies with examples specific to TCIM research is provided in Table 1. Additionally, the Concise, Logical, Explicit, Adaptive, and Reflective (CLEAR) Framework provides guidance for constructing high-quality prompts.[37] Concise prompts remove extraneous detail. Logical prompts follow a coherent sequence. Explicit prompts clarify expectations for content and structure. Adaptive prompting involves adjusting prompts based on task or discipline. Reflective prompting encourages researchers to iteratively review and refine GenAI chatbot outputs.

Because TCIM encompasses diverse practices and terminologies which may be underrepresented in LLM training data, domain-specific adaptation is necessary.[9,12,38] Researchers should define key terms, clarify concepts, and evaluate outputs critically. As GenAI chatbots become integrated into TCIM research workflows, researcher knowledge of prompt engineering will be essential to ensure that outputs support, rather than compromise, research quality and integrity.

4. Prompt engineering strategies for TCIM researchers

Using GenAI chatbots in TCIM research effectively requires more than posing questions; it involves the intentional design of prompts to ensure outputs are accurate, relevant, and culturally appropriate.[6,31,32,39–41]

4.1. Establishing context

Defining context is critical in TCIM prompts, as many modalities differ substantially from conventional biomedical paradigms. Prompts should clearly state the intervention, population, and research objective. For example, rather than asking for a general summary, specifying "Summarize this randomized controlled trial evaluating individualized homeopathic remedies for seasonal allergic rhinitis in adults" ensures the GenAI chatbot's output is more precise and aligned with the researcher's intent.

4.2. Structuring the prompt

Well-structured prompts improve clarity and coherence. Breaking requests into steps, such as outlining background sections or summarizing systematic reviews, can guide the GenAI chatbot to follow a logical format.[6,31,32,39–41] Including formatting preferences, such as bullet points or citation styles, can further align outputs with the formatting standards required for a given task.

4.3. Iterative refinement

Refinement is central to effective prompt engineering. Initial outputs may lack depth or contain inaccuracies, requiring follow-up prompts for clarification or alternative perspectives.[6,31,32,39–41] For instance, after receiving a broad explanation of "qi," a researcher might ask for analogies suited to a Western biomedical audience. Iteration supports both improved content and deeper understanding of GenAI chatbot limitations.

5. Use cases in health research

Within health research, GenAI chatbots can support a wide range of tasks,[4] but their use requires critical thinking and sound judgment. At early stages, they may assist with developing research questions, or drafting conceptual frameworks. In literature reviews, GenAI chatbots can help generate search terms, identify key concepts, and summarize abstracts. They may also introduce unfamiliar research methods or suggest reporting guidelines and appraisal tools, though researchers must ultimately assess their appropriateness. In quantitative research, chatbots can explain common statistical methods, suggest software options, or assist experienced users in drafting code. For qualitative research, they may support the creation of interview guides or suggest theoretical frameworks to inform coding and thematic analysis. GenAI chatbots can also help draft or adapt manuscripts, conference abstracts, and public-facing materials, including plain-language summaries or translations. Additionally, they may assist in writing, dissemination, and implementation activities across the research cycle.[1] In all cases, AI-generated content should be treated as a draft to be critically evaluated, verified, and refined before use. The following sections examine how these applications relate specifically to TCIM research.

6. Use cases in TCIM research

The following section describes GenAI chatbot use cases in TCIM research. Case examples and annotated prompts specific to TCIM research are provided in Table 2.

6.1. Research question development and study design

GenAI chatbots can be valuable tools for supporting the development of research questions and study designs, particularly in the early stages of TCIM research. They can assist researchers in brainstorming potential research topics, refining broad interests into focused, answerable questions, and exploring relevant study designs based on population, intervention, context, or outcomes.[3] For example, a GenAI chatbot might help a researcher transform a general curiosity about Ayurvedic treatments for insomnia into a specific, testable research question aligned with the Population, Intervention, Comparator, Outcome, Time (PICOT) framework.[42] Additionally, GenAI chatbots can suggest common methodologies, such as randomized controlled trials, observational studies, or qualitative approaches, and summarize their advantages and limitations. In fields like TCIM, where diverse epistemologies and intervention types may challenge conventional research paradigms, GenAI chatbots can prompt researchers to consider mixed methods designs or culturally appropriate frameworks. They may also aid in drafting logic models or conceptual frameworks to align study aims with theoretical underpinnings. Despite these benefits, researchers must apply critical thinking at every stage. GenAI chatbots can produce plausible-sounding suggestions that may be methodologically flawed, culturally inappropriate, or poorly aligned with the research context. Moreover, GenAI chatbots lack domain-specific judgment and may recommend designs that are infeasible or ethically problematic.[43-46]

6.2. Literature searches for background and narrative reviews

GenAI chatbots can serve as helpful assistants during the early stages of writing research manuscripts, particularly when conducting literature searches for background sections or narrative reviews. In this context, GenAI chatbots can be used to generate keyword suggestions, propose related concepts, and provide brief overviews of relevant topics.[47] For researchers working in TCIM, where terminology may vary widely across cultural contexts (e.g., "Ayurveda", "moxibustion", "Kampo"), these capabilities can help identify relevant sources that might otherwise be overlooked. GenAI chatbots may also be used to summarize abstracts or explain unfamiliar terms, thereby aiding researchers in developing a clearer understanding of the literature landscape.[47-49] However, GenAI chatbot-generated suggestions should always be treated as starting points, not definitive answers. GenAI chatbots can produce inaccurate or misleading information, omit key studies, or fail to grasp the cultural and contextual nuances,[50] critical to interpreting TCIM literature. In particular, studies have found that many widely-used GenAI chatbots have generated citations to sources which do not exist.[51,52] Researchers must apply critical thinking when assessing the relevance and credibility of GenAI chatbot outputs and should double check all findings using reliable sources such as peer-reviewed journal articles, textbooks, or clinical practice guidelines. Importantly, GenAI chatbots should be viewed as tools to support, not replace, human judgment. Used appropriately, they can accelerate the research process and spark new ideas, but researchers must remain responsible for evaluating and integrating the literature into their work with discernment and care.

6.3. Reviews with systematic search components

For structured reviews that involve a systematic search process, such as systematic or scoping reviews, GenAI chatbots can provide useful assistance in planning and preparation. GenAI chatbots may help formulate preliminary research questions, suggest relevant synonyms and Boolean operators, recommend appropriate databases, and guide users in incorporating controlled vocabularies like Medical Subject Headings (MeSH) or Emtree terms, for the Medical Literature Analysis and Retrieval System Online (MEDLINE) and Excerpta Medica database (EMBASE) bibliographic databases, respectively.[53] These features can be especially valuable for TCIM researchers, who often navigate a complex web of regional terminology and interdisciplinary subject matter. For instance, a GenAI chatbot might help a researcher explore equivalent or related terms for "herbal medicine" across various traditions, or generate initial ideas for classifying interventions in a review protocol.

Despite these benefits, at present, GenAI chatbots are not capable of conducting database searches, assessing study quality, or adhering to established methodological standards, nor are GenAI chatbots a replacement for review and development of search strategies by an information specialist.[50,53] They lack access to live indexing systems and may provide fabricated citations or incomplete information.[50] Researchers must therefore approach GenAI chatbot-generated content with caution and always verify outputs through manual searching and consultation with librarians or methodological experts. Critical thinking is essential in interpreting search strategies, determining inclusion criteria, and synthesizing results. Over-reliance on GenAI can compromise the integrity of the review,[29] particularly in a field like TCIM where context and cultural meaning greatly influence the interpretation of evidence. Ultimately, GenAI chatbots can play a valuable supporting role in the early phases of

Table 2
Case examples and annotated prompts for traditional, complementary, and integrative medicine research.

Section	Use Case Title	Initial Prompt	Refined Prompt	Outcome and Critical Reflection
6.1	Developing a Research Question on Ayurveda for Insomnia	"Help me create a research question about Ayurveda and sleep."	"Generate three research questions exploring the effectiveness of Ayurvedic herbal therapies for managing chronic insomnia in adults. Please structure them using the PICOT format (Population, Intervention, Comparator, Outcome, Time)."	This prompt yields clearer, well-framed research questions suitable for protocol development. The researcher might follow up by asking for rationale or design suggestions, but must verify alignment with evidence gaps and feasibility.
6.2	Searching Literature for a Narrative Review on Qigong	"Give me sources on qigong for cancer fatigue."	"Suggest key topics and concepts to include in a narrative review on the use of qigong for managing cancer-related fatigue. Summarize recent findings without providing citations, and focus on physiological mechanisms, patient-reported outcomes, and common limitations in the literature."	Thematic output supports review planning while reducing risk of hallucinated citations. Follow-up prompts can build on keyword development, but manual verification of sources remains necessary.
6.3	Designing a Search Strategy for a Scoping Review on Herbal Medicine	"Make a search strategy for herbal medicine."	"Help develop a preliminary search strategy for a scoping review on the use of herbal medicines in managing type 2 diabetes. Provide relevant keywords, MeSH terms, Boolean operators, and database recommendations. Present this in a format suitable for PubMed."	Results include a usable draft search string. The researcher should adapt this with the support of an information specialist to ensure comprehensiveness and reproducibility across databases.
6.4	Identifying Reporting Guidelines for a TCIM Trial	"What reporting guideline should I use for an acupuncture trial?"	"I am planning a randomized controlled trial on the effectiveness of acupuncture for generalized anxiety disorder. Please recommend a reporting guideline or extension suitable for non-pharmacologic treatments and briefly explain why it applies."	The revised prompt improves specificity, prompting accurate guideline suggestions and explanations. Researchers must still review guideline documents for proper application.
6.5	Writing R Code for Statistical Analysis in a Herbal Trial	"Give me R code for my trial."	"Please provide R code to conduct an independent samples *t*-test comparing changes in fasting blood glucose between treatment and control groups in a clinical trial of bitter melon for type 2 diabetes. Assume the data frame is called 'herbal_data' with variables 'group' and 'glucose_change'."	Code is functional and tailored. Researchers must verify its correctness, test it on sample data, and understand the statistical assumptions involved.
6.6	Generating Initial Themes for a Study on Mind-Body Therapies	"What themes come up in studies on yoga for pain?"	"Based on interview data from adults with chronic pain who use yoga, suggest initial themes and subthemes for a thematic analysis. Consider both physical and psychological dimensions, such as perceived pain relief, emotional regulation, self-efficacy, and spiritual growth."	Refined prompts produce richer thematic structures. Outputs may inspire coding frameworks but must be grounded in actual data and cultural context.
6.7	Drafting a Lay Summary on Ayurveda	"Summarize our Ayurveda paper in plain language."	"Draft a 200-word plain-language summary of a clinical study examining Ayurvedic herbs for anxiety. Write at an eighth-grade reading level, using simple vocabulary. Avoid making treatment claims and include a disclaimer recommending consultation with healthcare professionals."	The output improves in tone and clarity. Researchers should edit for accuracy, remove potential bias, and verify accessibility.
6.8	Brainstorming Implementation Strategies for a TCIM Intervention	"How can I implement this intervention in rural areas?"	"I am planning to implement a validated traditional Chinese medicine intervention for hypertension in rural primary care settings in India. Suggest implementation strategies using the CFIR framework, and identify potential barriers and facilitators related to provider training, patient acceptability, and infrastructure."	Tailored strategies align with implementation frameworks. Researchers must validate them through stakeholder input and site-specific planning.

systematic and scoping reviews, especially for brainstorming and generating preliminary strategies. However, the responsibility for conducting a transparent, reproducible, and rigorous review process remains firmly with the researcher.

6.4. Research methodology, reporting guidelines, and critical appraisal

While GenAI chatbots show potential to assist researchers with research methodology, reporting guidelines, and critical appraisal, there is currently a paucity of published evidence evaluating their effectiveness for these purposes. In practice, the most reliable contribution GenAI chatbots can offer in this domain is to help researchers identify potentially relevant methodological frameworks, reporting standards, and appraisal tools based on a study's aims, design, or topic area. GenAI chatbots can also help researchers, especially those with little or no prior knowledge of reporting guidelines, by identifying relevant guidelines and summarizing their purpose, scope, and methodological underpinnings in accessible terms. Some key examples include the Consolidated Standards of Reporting Trials (CONSORT) guideline for randomized controlled trials,[54] the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guideline for systematic reviews and meta-analyses,[55] and the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guideline for observational studies.[56] This function may be particularly helpful in TCIM, where interventions often involve complex, context-dependent, and culturally rooted elements, and require the use of TCIM-specific reporting guidelines or guideline extensions.[57,58] For example, a GenAI chatbot might suggest the CONSORT extension for non-pharmacologic treatments when planning a randomized controlled trial of acupuncture, or identify the PRISMA extension for a systematic review of herbal therapies.[59] Similarly, when seeking appraisal tools, a GenAI chatbot might point researchers to resources such as the Joanna Briggs Institute (JBI) Critical Appraisal Tools,[60] the Critical Appraisal Skills Programme's (CASP) checklists,[61] and the Oxford Centre for Evidence-Based Medicine's (CEBM) critical appraisal tools.[62] Researchers should nonetheless have some working knowledge of the reporting guidelines or critical appraisal tools they are using, in order to assess the relevance and accuracy of any GenAI chatbot-generated suggestions. GenAI

chatbots are not substitutes for critical thinking; their outputs should be treated as starting points for exploration, not definitive answers. To avoid misinformation or inappropriate application, researchers must critically evaluate all GenAI chatbot-generated content and verify it using reliable, authoritative sources.

6.5. Quantitative analysis

In quantitative TCIM research, GenAI chatbots may support researchers in identifying and learning about different statistical methods and software. For those with limited statistical background, GenAI chatbots can help by suggesting common analytic techniques, explaining basic concepts, or introducing software options that they may be unaware of, such as R, Python, SPSS, or STATA. This can be especially helpful in early stages of research, when a user is trying to select appropriate tools or learn about methods they were previously unaware of.[63] For example, a GenAI chatbot might describe when to use non-parametric tests like the Mann-Whitney U test or introduce generalized linear models for non-normal outcomes. For more statistically experienced researchers, GenAI chatbots may also assist in drafting or debugging code for specific analytical procedures. For instance, users might request help generating R or Python code to run chi-square tests, clean missing data in an acupuncture dataset, or visualize repeated measures data from an integrative medicine trial.[64-66] When used carefully, this support can reduce the time spent on routine tasks or troubleshooting syntax errors. However, current evidence suggests that caution is warranted when using GenAI chatbots for statistical analysis. Studies have shown that GenAI chatbot-generated code or statistical guidance can contain subtle errors, use inappropriate tests, or rely on flawed assumptions, especially if prompts lack sufficient context.[66-69] This makes GenAI outputs potentially misleading if accepted uncritically. Therefore, any GenAI chatbot-generated statistical content should be treated as a starting point, not a final product. All outputs must be reviewed line-by-line, tested on actual datasets, and validated against authoritative sources such as method papers, textbooks, or expert consultation.[65,66,69] Ultimately, while GenAI chatbots can assist in demystifying quantitative methods and improving workflow efficiency, final decisions about analysis and interpretation must rest with the research team. Critical thinking, domain-specific judgment, and a strong understanding of the data and research context remain indispensable.

6.6. Qualitative research

Qualitative methods are indispensable in TCIM research, offering insight into patient experiences, practitioner perspectives, and cultural narratives that are often not captured through quantitative approaches. GenAI chatbots can be leveraged to support various stages of qualitative research, particularly during early planning and design phases.[70] For example, researchers can prompt a GenAI chatbot to help develop semi-structured interview or focus group guides. By specifying the population (e.g., patients using herbal medicine for menopause), intervention (e.g., Ayurveda), and research aim (e.g., exploring treatment perceptions), users can generate tailored open-ended questions grounded in qualitative best practices.[42,71] In the data analysis phase, GenAI chatbots may be prompted to provide guidance on thematic analysis, grounded theory, or framework analysis.[72-74] For instance, a user can request examples of initial codes for a dataset on acupuncture use in chronic fatigue syndrome or ask for potential coding schemes based on known theoretical frameworks. However, while these tools may facilitate the logistics of qualitative work, they cannot replicate the interpretive and reflexive depth required for rigorous analysis. GenAI chatbot-generated coding suggestions or thematic groupings should never substitute for close reading of transcripts, deep immersion in the data, or engagement with the cultural and theoretical context of the research.[72,73,75–77] Moreover, there is a risk that GenAI chatbot-generated materials may inadvertently perpetuate stereotypes or overlook subtle forms of meaning.[73,78] Researchers must therefore critically assess GenAI chatbot contributions, drawing on their own expertise and ethical commitment to participant representation.

6.7. Writing and dissemination

GenAI chatbots can assist with the writing and dissemination of TCIM research findings across various platforms and audiences. Extreme caution should be taken when researchers use GenAI chatbots to help draft any portion of a manuscript for journal submission or an abstract for conference submission. While these tools can enhance efficiency and spark new ideas, researchers must apply critical thinking to assess the accuracy, appropriateness, and scientific rigour of any AI chatbot-generated content.[79] A cross-sectional audit of 162 scientific, technical, and medical publishers found that none allowed GenAI chatbots to be credited as authors on submitted manuscripts.[15] This policy reflects a core principle of scholarly publishing: authorship implies accountability, and GenAI chatbots are incapable of assuming ethical, legal, or intellectual responsibility for research outputs.[17,80] Researchers may also use GenAI chatbots to write cover letters, or generate lay summaries for public-facing materials, among other information relevant to scientific communication.[81-83] In addition, GenAI chatbots can help reframe study results for different audiences by adjusting tone, vocabulary, or structure,[84-86] for example, converting a technical description of a qigong intervention into a short plain-language summary suitable for a community health bulletin. Researchers working in multilingual or intercultural settings may also find GenAI chatbots helpful for drafting content in different languages or simplifying scientific language for public consumption.[87-89] Importantly, GenAI chatbots can support the creation of social media content to extend the reach of research findings. Researchers may prompt a GenAI chatbot to generate content for posts on platforms like Twitter/X, LinkedIn, or Instagram, such as threads, hashtags, or captions summarizing key findings of a TCIM study in engaging, accessible ways. This can be particularly valuable in resource-limited settings or for early-career researchers with limited science communication support. Social media content, in particular, poses unique risks given its brevity, visibility, and potential for misinterpretation.[90,91] Currently, there is limited scholarly literature offering guidance on how researchers can use GenAI chatbots to create social media content.

Regardless of the context, whenever GenAI chatbots are used to generate content intended for publication or public dissemination, their outputs should be treated as preliminary drafts rather than final versions. All content must be carefully reviewed, verified against the original research, and thoughtfully proofread and accordingly revised before sharing. Used responsibly, GenAI chatbots can streamline the writing and dissemination process.[92] However, researchers remain ultimately accountable for ensuring that all outputs, whether in journals, conference proceedings, press releases, or social media, are accurate, ethical, and contextually appropriate. Overreliance on GenAI chatbot-generated content without proper oversight can lead to the unintentional inclusion of factual errors, fabricated references, culturally inappropriate language, or otherwise misinterpretations of evidence.[30,82,93,94] Such oversights may undermine the credibility of the research, jeopardize publication, and erode trust in both the individual researcher and the broader scientific community.[95]

6.8. Implementation

Translating research findings into clinical, educational, or policy practice is essential to realizing the impact of TCIM research. Although implementation science in the context of AI has been explored across various health domains, there is limited evidence focused specifically on GenAI chatbots in implementation efforts.[96-98] GenAI chatbots can assist researchers in brainstorming strategies, identifying potential barriers and facilitators, and adapting interventions to specific settings.[99]

For example, a GenAI chatbot might help generate ideas for incorporating a traditional herbal therapy into primary care or tailoring a mindfulness intervention for rural health clinics. GenAI chatbots may also help draft logic models, develop implementation plans, or recommend frameworks such as the Consolidated Framework for Implementation Research (CFIR)[100,101] or the Reach, Efficacy, Adoption, Implementation, and Maintenance (RE-AIM) Model,[102,103] assuming the user has some familiarity with these tools. In educational contexts, GenAI chatbots can aid in preparing training materials or suggesting ways to integrate findings into professional development programs.[13,104] For policy implementation, GenAI chatbots can help draft preliminary summaries or frame messages aligned with health priorities.[105] Nevertheless, researchers must be cautious. GenAI chatbots do not possess contextual awareness of logistical, sociocultural, or resource-specific factors that are critical to implementation success.[99] Their outputs must be evaluated and refined through stakeholder consultation and expert input. While useful for early planning, GenAI chatbots cannot replace the nuanced, interdisciplinary expertise required to translate evidence into practice. Responsibility for customizing and executing implementation strategies ultimately rests with the researcher.

7. Limitations and challenges

GenAI chatbots show promise for supporting TCIM research, but their use comes with limitations. A key concern is the generation of hallucinations, responses that sound plausible but are factually incorrect or fabricated.[106,107] In health research, this may involve misrepresented findings, false citations, or inaccurate literature summaries.[26-28,93] For TCIM, where evidence can be limited or contested, such inaccuracies may inadvertently promote misinformation. Biases in GenAI chatbot training data also pose challenges. Most LLMs are developed using general internet and scientific text, which often reflect dominant biomedical paradigms.[108-111] As a result, TCIM practices may be misrepresented or dismissed. For instance, a GenAI chatbot may label acupuncture or herbal medicine as lacking evidence without considering the condition or context.[43-46] Ethical and epistemological issues also arise. Poorly designed prompts can reinforce stereotypes or oversimplify complex TCIM concepts.[47] Researchers should construct prompts that emphasize cultural sensitivity, accuracy, and nuance.[44-46] Output quality depends heavily on prompt clarity. Ambiguous or poorly structured prompts often yield vague or unhelpful results.[6,31,32,39–41] TCIM research often requires careful use of culturally specific terminology and interdisciplinary framing, which may be difficult for inexperienced users of GenAI chatbots.

Transparency is essential when using GenAI chatbots in scholarly work, in light of ethical and legal risks. Any AI-generated content that contributes to manuscripts or interpretations should be disclosed, in line with ethical publishing standards.[15,17,112] To date, numerous reporting guidelines have been developed to assist authors who use GenAI chatbots in their research across preclinical, translational, and clinical settings,[113,114] while others are undergoing development[115]; as such, it can be expected that TCIM-specific extensions may also be constructed in the future.[59] A major gap is the absence of strong regulatory frameworks governing AI use in medicine, including TCIM.[80,116,117] Issues such as informed consent, responsible disclosure, and safeguarding cultural knowledge require more attention. GenAI-assisted data analysis raises further concerns regarding privacy, especially when working with sensitive or identifiable information. Uploading such data to GenAI platforms may violate confidentiality and consent agreements, and even de-identified data could carry re-identification risks. Submitting data to third-party GenAI services may breach laws such as the General Data Protection Regulation (GDPR) or the Health Insurance Portability and Accountability Act (HIPAA).[80,118,119] This may also violate institutional review board (IRB) approvals, intellectual property agreements, or contractual terms.[120,121] Without clear oversight, the integrity of TCIM scholarship may be at risk.[43-46]

8. Conclusion

This guide highlights prompt engineering as a key skill for effectively using GenAI chatbots in TCIM research. Well-crafted prompts that are clear, structured, and culturally sensitive can improve the quality and relevance of AI-generated outputs across all stages of the TCIM research process. However, GenAI chatbots are not replacements for critical thinking or domain expertise. Their outputs should be treated as drafts, verified against authoritative sources. While these tools can improve efficiency, they also pose risks, such as the perpetuation of bias and misinformation. Responsible use requires ethical awareness, transparent reporting of GenAI chatbot involvement, and a commitment to scholarly and community standards. Thoughtful integration of GenAI chatbot use has the potential to strengthen scientific rigour and foster innovation across TCIM research.

CRediT authorship contribution statement

Jeremy Y. Ng: Conceptualization, Investigation, Methodology, Writing – original draft, Writing – review & editing.

Conflict of interest

The author is an editorial board member of the journal but his editorial board membership had no bearing on the decision of the article. The author has no conflicts of interest to declare.

Funding

None.

Ethics statement

This education article involved a review of the literature only; it did not require ethics approval or consent to participate.

Data availability

The data that support the findings of this study are available on request from the corresponding author.

References

1. Ng JY, Maduranayagam SG, Suthakar N, Li A, Lokker C, Iorio A, Haynes RB, Moher D. Attitudes and perceptions of medical researchers towards the use of artificial intelligence chatbots in the scientific process: an international cross-sectional survey. *Lancet Digit Health*. 2025;7:e94–102.
2. Kwon D. Is it OK for AI to write science papers? Nature survey shows researchers are split. *Nature*. 2025;641:574–578.
3. Kooli C. Chatbots in education and research: a critical examination of ethical implications and solutions. *Sustainability*. 2023;15:5614.
4. Khalifa M, Albadawy M. Using artificial intelligence in academic writing and research: an essential productivity tool. *Comput Methods Programs Biomed Update*. 2024:100145.
5. White J, Fu Q, Hays S, et al. *A prompt pattern catalog to enhance prompt engineering with ChatGPT*; 2023.
6. Giray L. Prompt engineering with ChatGPT: a guide for academic writers. *Ann Biomed Eng*. 2023;51:2629–2633.
7. Heston TF, Khun C. Prompt engineering in medical education. *Int Med Educ*. 2023;2:198–205.
8. Wang J, Shi E, Yu S, et al. *Prompt engineering for healthcare: methodologies and applications*; 2023.
9. Ng JY, Dhawan T, Dogadova E, Taghi-Zada Z, Vacca A, Wieland LS, Moher D. Operational definition of complementary, alternative, and integrative medicine derived from a systematic search. *BMC Compl Med Ther*. 2022;22:104.
10. Veziari Y, Leach MJ, Kumar S. Barriers to the conduct and application of research in complementary and alternative medicine: a systematic review. *BMC Compl Altern Med*. 2017;17:1660.
11. Ng JY, Cramer H, Lee MS, Moher D. Traditional, complementary, and integrative medicine and artificial intelligence: novel opportunities in healthcare. *Integr Med Res*. 2024;13:101024.
12. Raja M, Cramer H, Lee MS, Wieland LS, Ng JY. Addressing the challenges of traditional, complementary, and integrative medicine research: an international perspective and proposed strategies moving forward. *Perspect Integr Med*. 2024;3:86–97.

13. Ng JY, Shah AQ, Roni E, et al. Attitudes and perceptions of university students and postdoctoral fellows in the medical and life sciences towards the use of artificial intelligence chatbots in the educational process: a large-scale, international cross-sectional survey. *medRxiv. Published*. 2025;26. doi:10.1101/2025.06.25.25328523.
14. Ng JY, Krishnamurthy M, Deol G, et al. Attitudes and perceptions of biomedical journal editors-in-chief towards the use of artificial intelligence chatbots in the scholarly publishing process: a cross-sectional survey. *medRxiv. Published May*. 2025;27. doi:10.1101/2025.05.26.25328101.
15. Bhavsar D, Duffy L, Jo H, Lokker C, Haynes RB, Iorio A, Marusic A, Ng JY. Policies on artificial intelligence chatbots among academic publishers: a cross-sectional audit. *Res Integr Peer Rev*. 2025;10:1.
16. International Association of Scientific. *Recommendations For a Classification of AI Use in Academic Manuscript Preparation*. Technical & Medical Publishers; June 25, 2025 STM Website, *https://stm-assoc.org/document/recommendations-for-a-classification-of-ai-use-in-academic-manuscript-preparation/* .
17. International Committee of Medical Journal Editors. Defining the role of authors and contributors. ICMJE Website. https://www.icmje.org/recommendations/browse/roles-and-responsibilities/defining-the-role-of-authors-and-contributors.html. Published 2025. Accessed June 25, 2025.
18. Zielinski C, Winker MA, Aggarwal R, Ferris LE, Heinemann M, Jr Lapeña JF, Pai SA, Citrome L, Alam M, Voight M, Habibzadeh F. Chatbots, generative AI, and scholarly manuscripts: WAME recommendations on chatbots and generative artificial intelligence in relation to scholarly publications. *Colomb Med*. 2023;54:e106.
19. . *Authorship and AI Tools*. Committee on Publication Ethics; 2025 COPE Website *https://publicationethics.org/guidance/cope-position/authorship-and-ai-tools* Published 2025. Accessed June 25.
20. European Association of Science Editors. Recommendations on the use of AI in scholarly communication. EASE Website. https://ease.org.uk/2024/09/recommendations-on-the-use-of-ai-in-scholarly-communication/. Published 2024. Accessed June 25, 2025.
21. International Business Machines. What are large language models (LLMs)? IBM Website. https://www.ibm.com/think/topics/large-language-models. Published 2025. Accessed June 25, 2025.
22. OpenAI Platform. Optimizing LLM accuracy. OpenAI Website. https://platform.openai.com/docs/guides/optimizing-llm-accuracy. Published 2025. Accessed June 25, 2025.
23. Google AI for DevelopersPrompt design strategies. *Google website*. 2025. https://ai.google.dev/gemini-api/docs/prompting-strategies. Published 2025. Accessed June 25.
24. Singh V. Exploring the role of large language model (LLM)-based chatbots for human resources [dissertation]. *Austin: The University of Texas at Austin*. 2023. https://repositories.lib.utexas.edu/server/api/core/bitstreams/36eac695-555e-4a00-9135-4a543ac46469/content.
25. Dam SK, Hong CS, Qiao Y, et al. *A complete survey on LLM-based AI chatbots*; 2024.
26. Walters WH, Wilder EI. Fabrication and errors in the bibliographic citations generated by ChatGPT. *Sci Rep*. 2023;13:14045.
27. Chen A, Chen DO. Accuracy of chatbots in citing journal articles. *JAMA Netw Open*. 2023;6:e2327647.
28. Graf EM, McKinney JA, Dye AB, Lin L, Sanchez-Ramos L. Exploring the limits of artificial intelligence for referencing scientific articles. *Am J Perinatol*. 2024;41:2072–2081.
29. Kasani PH, Cho KH, Jang JW, Yun CH. Influence of artificial intelligence and chatbots on research integrity and publication ethics. *Sci Ed*. 2024;11:12–25.
30. Dwivedi YK, Kshetri N, Hughes L, Slade EL, Jeyaraj A, Kar AK, Baabdullah AM, Koohang A, Raghavan V, Ahuja M, Albanna H. Opinion paper: "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. *Int J Inf Manag*. 2023;71:102642.
31. International Business Machines. What is prompt engineering? IBM Website. https://www.ibm.com/think/topics/prompt-engineering. Published 2025. Accessed June 25, 2025.
32. OpenAI Platform. Text generation and prompting. OpenAI Website. https://platform.openai.com/docs/guides/text?api-mode=chat. Published 2025. Accessed June 25, 2025.
33. International Business Machines. What is zero-shot prompting? IBM Website. https://www.ibm.com/think/topics/zero-shot-prompting. Published 2025. Accessed June 25, 2025.
34. International Business Machines. What is chain of thought (CoT) prompting? IBM Website. https://www.ibm.com/think/topics/chain-of-thoughts. Published 2025. Accessed June 25, 2025.
35. Sun S, Zhuang S, Wang S, Zuccon G. An investigation of prompt variations for zero-shot LLM-based rankers. *European Conference on Information Retrieval*. Cham: Springer Nature Switzerland; 2025:185–201.
36. International Business Machines. Use role prompting with IBM Watsonx and Granite. IBM Website. https://www.ibm.com/think/tutorials/using-role-prompting-with-watsonx-and-granite. Published 2025. Accessed June 25, 2025.
37. Lo LS. The CLEAR path: a framework for enhancing information literacy through prompt engineering. *J Acad Librariansh*. 2023;49:102720.
38. Ng JY, Dhawan T, Fajardo RG, Masood HA, Sunderji S, Wieland LS, Moher D. The brief history of complementary, alternative, and integrative medicine terminology and the development and creation of an operational definition. *Integr Med Res*. 2023;12:100978.
39. Ekin S. Prompt engineering for ChatGPT: a quick guide to techniques, tips, and best practices. Authorea Preprints. *TechRxiv*. 2023;30.
40. ChatGPT Ray PP. A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. *Internet Things Cyber-Phys Syst*. 2023;3:121–154.
41. Sarkar R, Sarrafzadeh B, Chandrasekaran N, et al. *Conversational user-AI intervention: a study on prompt rewriting for improved LLM response generation*; 2025.
42. Gosak L, Štiglic G, Pruinelli L, Vrbnjak D. PICOT questions and search strategies formulation: a novel approach using artificial intelligence automation. *J Nurs Scholarsh*. 2025;57:5–16.
43. Ng JY, Dimitriou I, Pesala B, Kajiwara M, Pillai GK. From ancient wisdom to modern practices: transformative potential of digital health innovations in advancing traditional medicine. *J Complement Integr Med*. 2025;22:209–212.
44. Amabie T, Izah SC, Ogwu MC, Hait M. Harmonizing tradition and technology: the synergy of artificial intelligence in traditional medicine. *Herbal Medicine phytochemistry: Applications and Trends*. Cham: Springer International Publishing; 2023:1–23.
45. Ofosu-Asare Y. Cognitive imperialism in artificial intelligence: counteracting bias with indigenous epistemologies. *AI Soc*. 2025;40:3045–3061.
46. World Health Organization. Mapping the application of artificial intelligence in traditional medicine: technical brief. WHO Website. https://www.who.int/publications/i/item/9789240107663. Published 2025. Accessed June 25, 2025.
47. Lund BD, Wang T, Mannuru NR, Nie B, Shimray S, Wang Z. ChatGPT and a new academic reality: artificial intelligence-written research papers and the ethics of the large language models in scholarly publishing. *J Assoc Inf Sci Technol*. 2023;74:570–581.
48. Glickman M, Zhang Y AI. and generative AI for research discovery and summarization. *Harv Data Sci Rev*. 2024;6(2). doi:10.1162/99608f92.7f9220ff.
49. Hwang T, Aggarwal N, Khan PZ, Roberts T, Mahmood A, Griffiths MM, Parsons N, Khan S. an ChatGPT assist authors with abstract writing in medical journals? Evaluating the quality of scientific abstracts generated by ChatGPT and original abstracts. *PLoS One*. 2024;19:e0297701.
50. Gwon YN, Kim JH, Chung HS, Jung EJ, Chun J, Lee S, Shim SR. The use of generative AI for scientific literature searches for systematic reviews: chatGPT and Microsoft Bing AI performance evaluation. *JMIR Med Inform*. 2024;12:e51187.
51. Aljamaan F, Temsah MH, Altamimi I, Al-Eyadhy A, Jamal A, Alhasan K, Mesallam TA, Farahat M, Malki KH. Reference hallucination score for medical artificial intelligence chatbots: development and usability study. *JMIR Med Inform*. 2024;12:e54345.
52. Chelli M, Descamps J, Lavoué V, Trojani C, Azar M, Deckert M, Raynier JL, Clowez G, Boileau P. Ruetsch-Chelli C. Hallucination rates and reference accuracy of ChatGPT and bard for systematic reviews: comparative analysis. *J Med Internet Res*. 2024;26:e53164.
53. Park SG, Carroll M, Esteve LM, Singh K. *Exploring generative AI and natural language processing to develop search strategies for systematic reviews*. Portland, OR: American Society for Engineering Education; 2024.
54. Schulz KF, Altman DG, Moher D, Group CONSORT. CONSORT 2010 statement: updated guidelines for reporting parallel group randomized trials. *Ann Intern Med*. 2010;152:726–732.
55. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, Shamseer L, Tetzlaff JM, Akl EA, Brennan SE, Chou R. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*. 2021:372.
56. Von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Lancet*. 2007;370:1453–1457.
57. Ng JY, Wieland LS, Lee MS, Liu JP, Witt CM, Moher D, Cramer H. Open science practices in traditional, complementary, and integrative medicine research: a path to enhanced transparency and collaboration. *Integr Med Res*. 2024;13:101047.
58. Ng JY, Lee MS, Liu JP, Steel A, Wieland LS, Witt CM, Moher D, Cramer H. How can meta-research be used to evaluate and improve the quality of research in the field of traditional, complementary, and integrative medicine? *Integr Med Res*. 2024;13:101068.
59. Enhancing the Quality and Transparency Of Health Research Network. Search for "complementary and alternative medicine" reporting guidelines. EQUATOR Network Website. https://www.equator-network.org/reporting-guidelines-medical-specialty/complementary-and-alternative-medicine/?post_type=eq_guidelines&eq_guidelines_study_design=0. Published 2025. Accessed June 25, 2025.
60. Joanna Briggs Institute. Critical appraisal tools. JBI Website. https://jbi.global/critical-appraisal-tools. Published 2025. Accessed June 25, 2025.
61. Critical Appraisal Skills Programme. CASP checklists. CASP Website. https://casp-uk.net/casp-tools-checklists/. Published 2025. Accessed June 25, 2025.
62. Centre for Evidence-Based Medicine. Critical appraisal tools. CEBM Website. https://www.cebm.ox.ac.uk/resources/ebm-tools/critical-appraisal-tools. Published 2025. Accessed June 25, 2025.
63. OpenAI. Data analysis with ChatGPT. OpenAI Website. https://help.openai.com/en/articles/8437071-data-analysis-with-chatgpt. Published 2025. Accessed June 25, 2025.
64. Ruta MR, Gaidici T, Irwin C, Lifshitz J. ChatGPT for univariate statistics: validation of AI-assisted data analysis in healthcare research. *J Med Internet Res*. 2025;27:e63550.
65. Prandner D, Wetzelhütter D, Hese S. ChatGPT as a data analyst: an exploratory study on AI-supported quantitative data analysis in empirical research. *Front Educ*. 2025;9:1417900.
66. Huang Y, Wu R, He J, Xiang Y. Evaluating ChatGPT-4.0's data analytic proficiency in epidemiological studies: a comparative analysis with SAS, SPSS, and R. *J Glob Health*. 2024;14:04070.

67. Ordak M. ChatGPT's skills in statistical analysis using the example of allergology: do we have reason for concern? *Healthcare*. 2023;11:2554.
68. Koçak D. Examination of ChatGPT's performance as a data analysis tool. *Educ Psychol Meas*. 2025. doi:10.1177/00131644241302721.
69. Schwarz J. The use of generative AI in statistical data analysis and its impact on teaching statistics at universities of applied sciences. *Teach Stat*. 2025.
70. Zhang H, Wu C, Xie J, Lyu Y, Cai J, Carroll JM. Harnessing the power of AI in qualitative research: exploring, using and redesigning ChatGPT. *Comput Hum Behav Artif Humans*. 2025;4:100144.
71. Christou P. The use of artificial intelligence (AI) in qualitative research for theory development. *Qual Rep*. 2023;28:2739–2755.
72. Wachinger J, Bärnighausen K, Schäfer LN, Scott K, McMahon SA. Prompts, pearls, imperfections: comparing ChatGPT and a human researcher in qualitative data analysis. *Qual Health Res*. 2024 May 22:10497323241244669.
73. Lee VV, van der Lubbe SC, Goh LH, Valderas JM. Harnessing ChatGPT for thematic analysis: are we ready? *J Med Internet Res*. 2024;26:e54974.
74. Turobov A., Coyle D., Harding V. Using ChatGPT For Analytics. Bennett Institute for Public Policy, University of Cambridge Website. https://www.bennettinstitute.cam.ac.uk/wp-content/uploads/2024/05/Using-ChatGPT-for-analytics-WP.pdf. Published May 2024. Accessed June 25, 2025.
75. Nguyen-Trung K. ChatGPT in thematic analysis: can AI become a research assistant in qualitative research? *Qual Quant*. 2025 Jun 2:1–34.
76. Bijker R, Merkouris SS, Dowling NA, Rodda SN. ChatGPT for automated qualitative research: content analysis. *J Med Internet Res*. 2024;26:e59050.
77. Morgan DL. Exploring the use of artificial intelligence for qualitative data analysis: the case of ChatGPT. *Int J Qual Methods*. 2023;22:16094069231211248.
78. Sakaguchi K, Sakama R, Watari T. Evaluating ChatGPT in qualitative thematic analysis with human researchers in the Japanese clinical context and its cultural interpretation challenges: comparative qualitative study. *J Med Internet Res*. 2025;27:e71521.
79. Huang J, Tan M. The role of ChatGPT in scientific communication: writing better scientific review articles. *Am J Cancer Res*. 2023;13:1148.
80. Bird E, Fox-Skelly J, Jenner N, Larbey R, Weitkamp E, Winfield A. *The Ethics of Artificial intelligence: Issues and Initiatives*. European Parliamentary Research Service; 2020.
81. Pollesello P, Papp Z. Integrating an artificial intelligence chatbot in scientific communication: dos and don'ts. *Eur Sci Ed*. 2023;49:e112023.
82. Hendriks F, Barel-Ben David Y, Banse L, Fick J, Greussing E, Klein-Avraham I, Rakedzon T, Taddicken M, Baram-Tsabari A. Generative AI in science communication: fostering scientists' good working habits for ethical and effective use. *Sci Commun*. 2025:10755470251343486.
83. Valle D, Barone M. *Generative AI For Science Communication*. Fermi National Accelerator Laboratory Website; 2025 *https://inspirehep.net/files/1dee4deb1cdfbe30c922b67d3eeeea39* Published August 2024. Accessed June 25.
84. Li G., Alshomary M., Muresan S. "Is ChatGPT a better explainer than my professor?": evaluating the explanation capabilities of LLMs in conversation compared to a human baseline. arXiv preprint arXiv:2406.18512. 2024 Jun 26.
85. Markowitz DM. From complexity to clarity: how AI enhances perceptions of scientists and the public's understanding of science. *PNAS Nexus*. 2024;3(9):pgae387.
86. Faris M. Use of AI in Communication and Dissemination Statistics. UNECE Website. https://unece.org/sites/default/files/202310/DissComm2023_S3_Indonesia_Faris_Paper.pdf. Published October 2023. Accessed June 25, 2025.
87. Akter SN, Yu Z, Muhamed A, et al. *Neubig G. An in-depth look at Gemini's language abilities*; 2023.
88. Orlando M., Liao S., Kruger J.L. Translation and Interpreting technologies and Their Impact On the Industry. Macquarie University Website. https://researchers.mq.edu.au/en/publications/translation-and-interpreting-technologies-and-their-impact-on-the. Published February 2024. Accessed June 25, 2025.
89. Ahn S. The transformative impact of large language models on medical writing and publishing: current applications, challenges and future directions. *Korean J Physiol Pharmacol*. 2024;28(5):393–401.
90. Tandar CE, Bajaj SS, Stanford FC. Social media and artificial intelligence—Understanding medical misinformation through Snapchat's new artificial intelligence chatbot. *Mayo Clin Proc Digit Health*. 2024;2(2):252–254.
91. Morita PP, Lotto M, Kaur J, Chumachenko D, Oetomo A, Espiritu KD, Hussain IZ. What is the impact of artificial intelligence-based chatbots on infodemic management? *Front Public Health*. 2024;12:1310437.
92. Resnik DB, Hosseini M. The ethics of using artificial intelligence in scientific research: new guidance needed for a new tool. *AI Ethics*. 2024:1–23.
93. Loh E. ChatGPT and generative AI chatbots: challenges and opportunities for science, medicine and medical leaders. *BMJ Lead*. 2023;8(1):e000797.
94. Kacena MA, Plotkin LI, Fehrenbacher JC. The use of artificial intelligence in writing scientific review articles. *Curr Osteoporos Rep*. 2024;22(1):115–121.
95. Williamson SM, Prybutok V. The era of artificial intelligence deception: unraveling the complexities of false realities and emerging threats of misinformation. *Information*. 2024;15(6):299.
96. Hogg HD, Al-Zubaidy M, Keane PA, Hughes G, Beyer FR, Maniatopoulos G. Evaluating the translation of implementation science to clinical artificial intelligence: a bibliometric study of qualitative research. *Front Health Serv*. 2023;3:1161822.
97. Gama F, Tyskbo D, Nygren J, Barlow J, Reed J, Svedberg P. Implementation frameworks for artificial intelligence translation into health care practice: scoping review. *J Med Internet Res*. 2022;24(1):e32215.
98. Nilsen P, Svedberg P, Nygren J, Frideros M, Johansson J, Schueller S. Accelerating the impact of artificial intelligence in mental healthcare through implementation science. *Implement Res Pract*. 2022;3:26334895221112033.
99. Trinkley KE, An R, Maw AM, Glasgow RE, Brownson RC. Leveraging artificial intelligence to advance implementation science: potential opportunities and cautions. *Implement Sci*. 2024;19(1):17.
100. Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. *Implement Sci*. 2009;4:50.
101. Damschroder LJ, Reardon CM, Widerquist MA, Lowery J. The updated Consolidated Framework for Implementation Research based on user feedback. *Implement Sci*. 2022;17(1):75.
102. Glasgow RE, Vogt TM, Boles SM. Evaluating the public health impact of health promotion interventions: the RE-AIM framework. *Am J Public Health*. 1999;89(9):1322–1327.
103. Glasgow RE, Harden SM, Gaglio B, Rabin B, Smith ML, Porter GC, Ory MG, Estabrooks PA. RE-AIM planning and evaluation framework: adapting to new science and practice with a 20-year review. *Front Public Health*. 2019;7:64.
104. Singh B, Olds T, Brinsley J, Dumuid D, Virgara R, Matricciani L, Watson A, Szeto K, Eglitis E, Miatke A, Simpson CE. Systematic review and meta-analysis of the effectiveness of chatbots on lifestyle behaviours. *NPJ Digit Med*. 2023;6(1):118.
105. Reddy S. Generative AI in healthcare: an implementation science informed translational path on application, integration and governance. *Implement Sci*. 2024;19(1):27.
106. Huang L, Yu W, Ma W, Zhong W, Feng Z, Wang H, Chen Q, Peng W, Feng X, Qin B, Liu T. A survey on hallucination in large language models: principles, taxonomy, challenges, and open questions. *ACM Trans Inf Syst*. 2025;43(2):1–55.
107. Teubner T, Flath CM, Weinhardt C, Van Der Aalst W, Hinz O. Welcome to the era of ChatGPT et al. the prospects of large language models. *Bus Inf Syst Eng*. 2023;65(2):95–101.
108. Cirillo D, Catuara-Solarz S, Morey C, Guney E, Subirats L, Mellino S, Gigante A, Valencia A, Rementeria MJ, Chadha AS, Mavridis N. Sex and gender differences and biases in artificial intelligence for biomedicine and healthcare. *NPJ Digit Med*. 2020;3(1):81.
109. Chow JC, Li K. Ethical considerations in human-centered AI: advancing oncology chatbots through large language models. *JMIR Bioinform Biotechnol*. 2024;5:e64406.
110. Nazi ZA, Peng W. Large language models in healthcare and medical domain: a review. *Informatics. Basel* MDPI; 2024:57.
111. Fletcher RR, Nakeshimana A, Olubeko O. Addressing fairness, bias, and appropriate use of artificial intelligence and machine learning in global health. *Front Artif Intell*. 2021;3:561802.
112. Lund BD, Naheem KT. Can ChatGPT be an author? A study of artificial intelligence authorship policies in top academic journals. *Learn Publ*. 2024;37(1):13–21.
113. Kolbinger FR, Veldhuizen GP, Zhu J, Truhn D, Kather JN. Reporting guidelines in medical artificial intelligence: a systematic review and meta-analysis. *Commun Med*. 2024;4(1):71.
114. Enhancing the Quality and Transparency Of Health Research Network. Search for "artificial intelligence/machine learning studies" reporting guidelines. EQUATOR Network Website. https://www.equator-network.org/?post_type=eq_guidelines&eq_guidelines_study_design=artificial-intelligence-machine-learning-studies&eq_guidelines_clinical_specialty=0&eq_guidelines_report_section=0&s=&btn_submit=Search+Reporting+Guidelines. Published 2025. Accessed August 7, 2025.
115. Enhancing the Quality and Transparency Of Health Research Network. Reporting guidelines under development for other study designs: artificial intelligence and machine learning, modeling, coding, informatics. EQUATOR Network Website. https://www.equator-network.org/library/reporting-guidelines-under-development/reporting-guidelines-under-development-for-other-study-designs/#AIMOD. Published 2025. Accessed August 7, 2025.
116. Minssen T, Vayena E, Cohen IG. The challenges for regulating medical use of ChatGPT and other large language models. *JAMA*. 2023;330(4):315–316.
117. Fecher B, Hebing M, Laufer M, Pohle J, Sofsky F. riend or foe? Exploring the implications of large language models on the science system. *AI Soc*. 2025;40(2):447–459.
118. Khalid N, Qayyum A, Bilal M, Al-Fuqaha A, Qadir J. Privacy-preserving artificial intelligence in healthcare: techniques and applications. *Comput Biol Med*. 2023;158:106848.
119. Rezaeikhonakdar D. AI chatbots and challenges of HIPAA compliance for AI developers and vendors. *J Law Med Ethics*. 2023;51(4):988–995.
120. Pantanowitz L, Hanna M, Pantanowitz J, Lennerz J, Henricks WH, Shen P, Quinn B, Bennet S, Rashidi HH. Regulatory aspects of AI-ML. *Mod Pathol*. 2024:100609.
121. U.S. Department of Health and Human Services. IRB considerations on the use of artificial intelligence in human subjects research. HHS Website. https://www.hhs.gov/ohrp/sachrp-committee/recommendations/irb-considerations-use-artificial-intelligence-human-subjects-research/index.html. Published 2025. Accessed June 25, 2025.