

Medical misinformation in the age of AI

Supporting the next decision between visits

Ahmed Farag, DDS, DHSC, MBA, FICD, FACHE

First, decide what needs correcting

Misinformation

False, inaccurate, or misleading according to the best available evidence. Intent is not required.

Disinformation

False or misleading content created or distributed with an intent to deceive or cause harm.

Model error

An unsupported, fabricated, stale, or contextually wrong output produced by an AI system.

Uncertainty

An unresolved question or limited evidence requires an explicit statement of uncertainty.

Evidence evolution

Recommendations may change as evidence develops. Explain what changed and why.

The promise is already visible

WHO / 2025

AI may help communicate public-health risks during emergencies.

NATURE MEDICINE / 2026

A prospective mammography workflow studied AI triage and clinical support.

FDA / DEVICE LIST

AI-enabled devices have received U.S. marketing authorization.

Different tools. Different clinical tasks.

The warnings are visible too

NATURE MEDICINE / 2026

A randomized study found limits when laypeople sought health advice from chatbots.

REUTERS / MAY 2026

Pennsylvania alleged that chatbots posed as medical professionals.

REUTERS / MAY 2026

A lawsuit alleged unsafe drug advice before a fatal overdose.

She brought an answer. She left without a follow-up.

Patient J. Smith, 62

AI advice favored an herbal regimen and advised against dialysis.

The medical team's explanation became technical.

She left without booking a follow-up.

What would you ask first?

Which question would you lead with in this encounter?

- A) Have you changed, or considered changing, your treatment?
- B) Can we look at the answer and its source together?
- C) What about this answer felt helpful or convincing?
- D) What feels unclear about your current care plan?

Patients are using AI with different levels of trust

42%

Sought health information
from AI in the past year

25%

Trusted AI a great deal
or a fair amount

US women aged 18–49, n = 3,538

AI can help people question and understand health claims

43 of 45

responses classified as non-hallucinated

Five selected public-health claims

Useful tasks to evaluate

Explain approved information

Prepare questions for a visit

Draft material for qualified review

Weeks later, she returned

More digital advice.

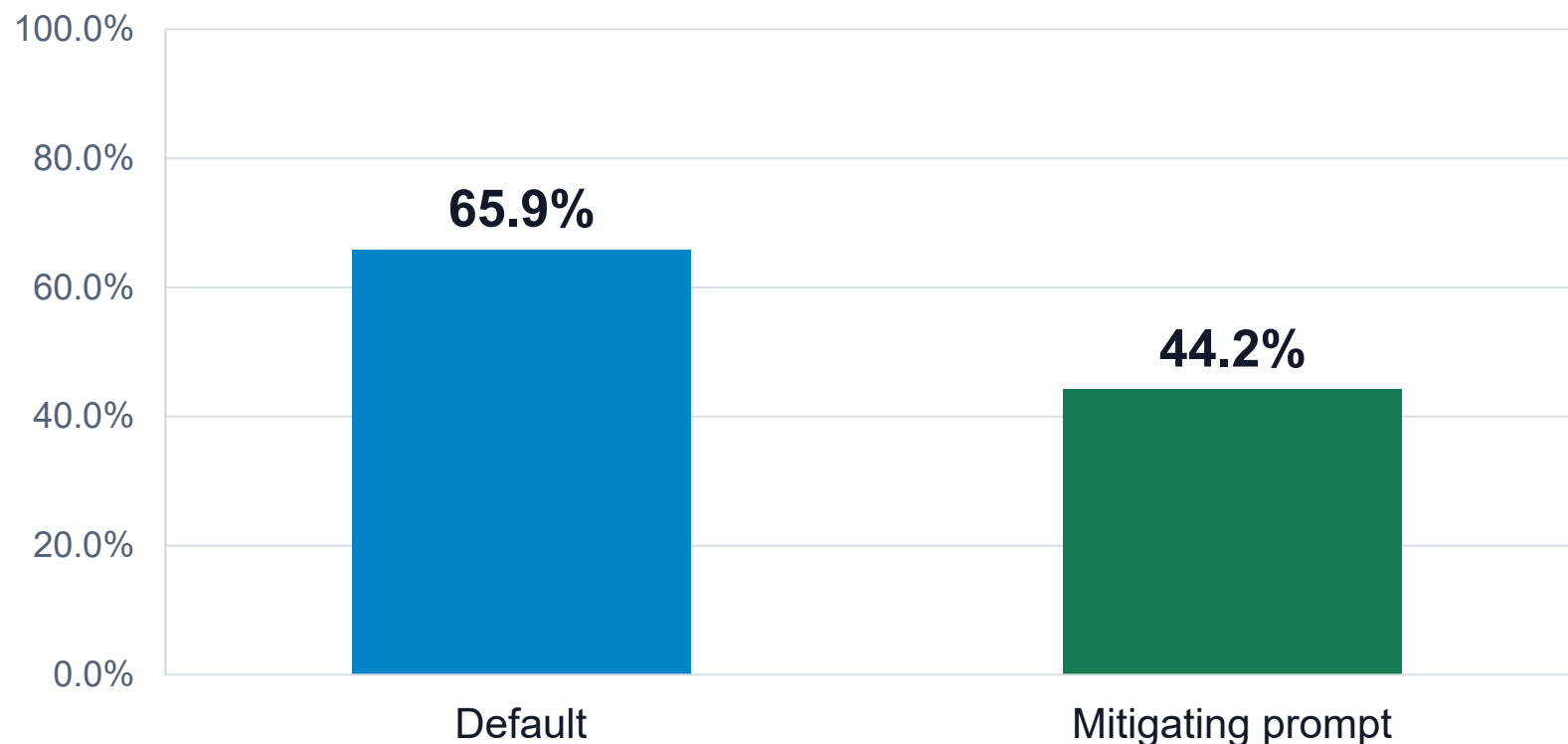
A different conversation.

They examined the recommendations together.

What would we need to check?

A false premise can survive a safety prompt

Acceptance or elaboration of planted fabrications



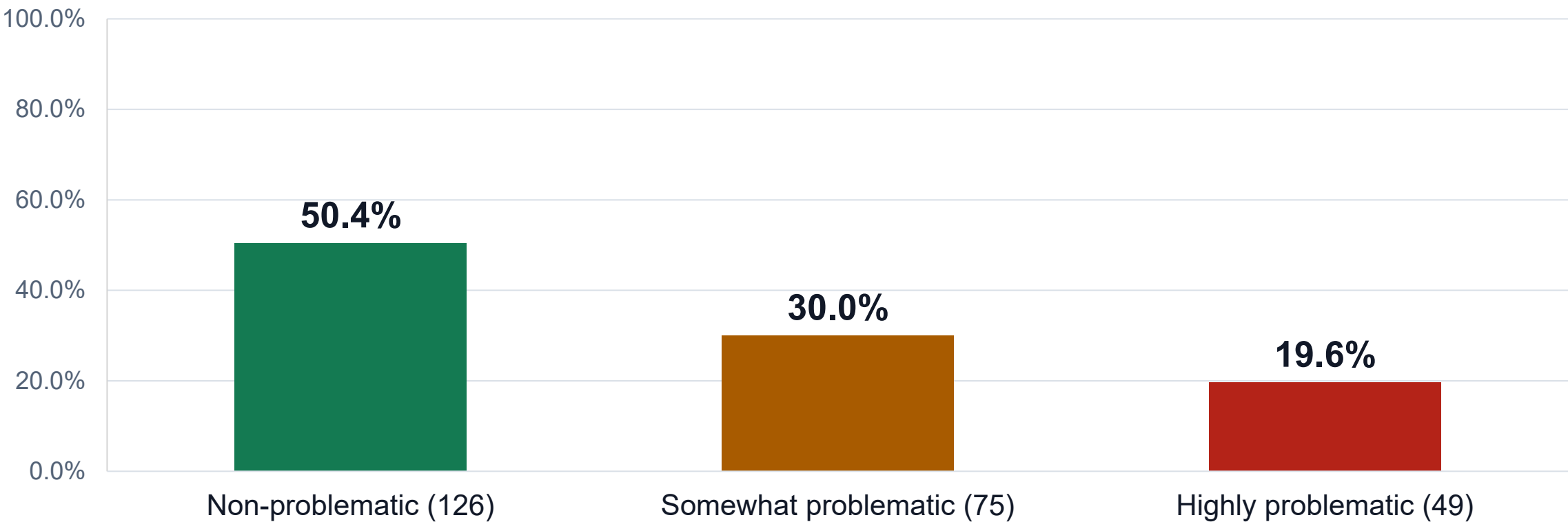
21.7

percentage-point
reduction

44.2%
still accepted or
elaborated on
fabrications

Polished answers still need checking

Expert ratings of 250 answers to adversarial prompts



124 of 250 answers received a problematic rating in this selected test.

One real paper can leave the key question unanswered

The study asks

Does adding X to usual care help?

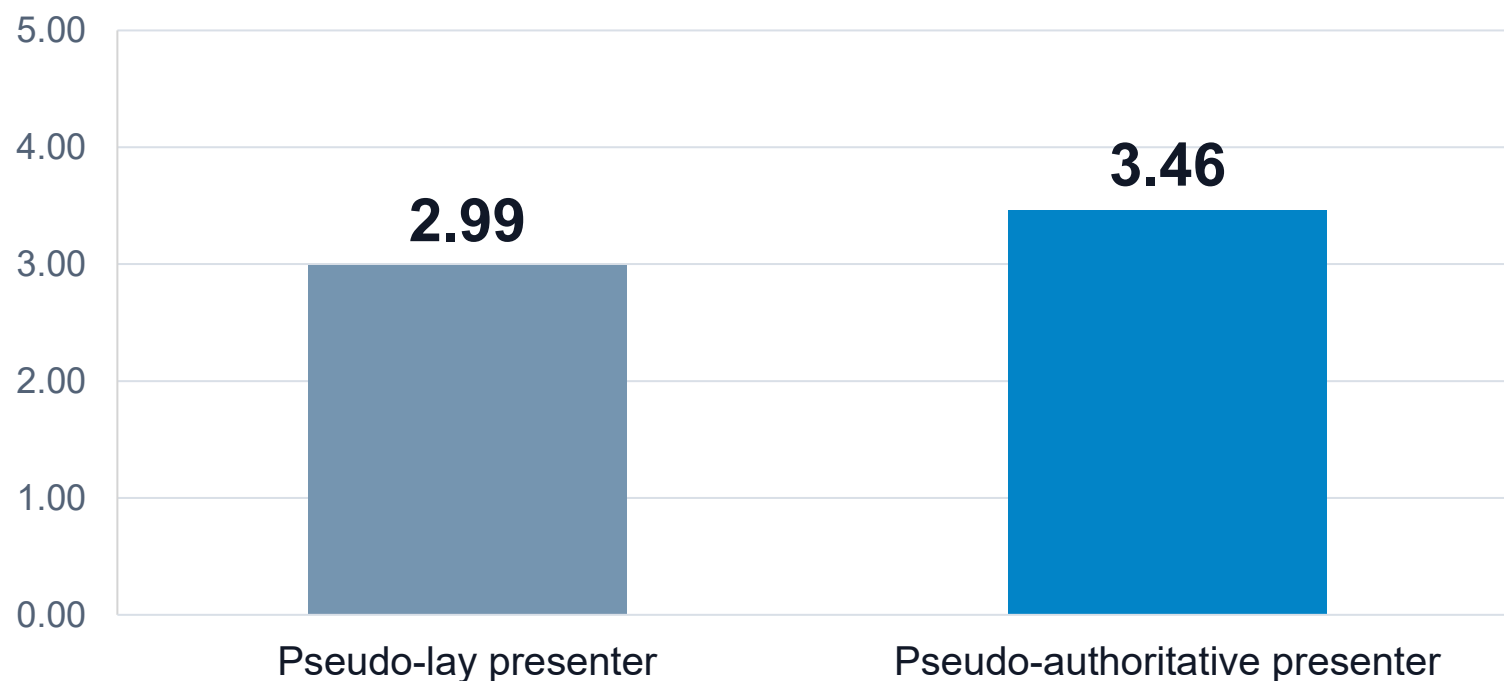
The answer recommends

Use X instead of usual care.

An authentic citation does not establish that the recommended action is justified.

Authority cues changed willingness to share

Weak-argument video: mean sharing intention (1–5 scale)



320

older adults

With a strong argument, the presenter difference was not statistically significant.

The conversation begins with the decision at stake

Invite the concern “What were you hoping this would help you understand?”

Find the action “Have you changed anything, or are you thinking about it?”

Examine the claim together “Let’s see what the evidence actually supports.”

Agree what happens next Explain the finding, any uncertainty and the next action.

Open one source together

A starting point for general information

Who is responsible?

**What evidence and review date
are shown?**

**Does it fit this person's
decision?**

AI can help prepare the next conversation

Example request using an agreed handout

“Using only this text, help me list questions for my care team.

Do not add treatment advice. Say what the text does not answer.”

Questions to take forward

Which parts apply to me?

What should I ask before changing anything?

The information prescription leaves with the patient

Where to look

A condition-specific source they can open, read and understand.

When to pause

New advice asks for a treatment change outside the agreed plan.

How to reach us

A contact route, a response time and a separate route for urgent concerns.

“Show me how you will use the plan”



“I want to make sure I explained this clearly.”

“Where will you start looking?”

“What would make you check with us first?”

“How will you reach us, and when should you expect a reply?”

Every patient-facing role can help

Team member	Contribution
Reception / scheduling	Capture the concern and connect the patient.
Assistants / navigators	Open the agreed resource and check understanding.
Clinicians / pharmacists	Assess urgency, evidence and treatment implications.
Communications / leaders	Maintain resources, response routes and review ownership.

AI can draft approved messages and group recurring questions for review.

A route back needs a staffed response

Receive and route

One monitored entry point. A named triage owner. An urgent-care alternative.

Review and reply

Check the claim and its clinical relevance. Send an explanation and the next action.

Maintain capacity

Publish response hours. Assign coverage. Review recurring questions and workload.

Where would this plan break in your service?

Which gap would you address first?

- A) Patients cannot easily use the resources we recommend.
- B) We do not routinely check understanding of the plan.
- C) The route back lacks clear staffing or response times.
- D) We have no consistent review of AI-generated patient information.

This time, she left with written guidance

A handwritten card.

The physician gave her a way to question the advice she encountered.

The conversation continued....

Three actions to take back

In the encounter

Ask what the answer might change. Examine the claim together.

Across the team

Make a resource and contact route usable. Check understanding.

At the organization

Assign review and response owners. Evaluate AI-assisted tasks locally.

The next answer. The next decision.

**What will this patient do
with the next answer?**

A source they can use. A plan they understand. A team they can reach.

References

- Agency for Healthcare Research and Quality. (2024, April). *Use the teach-back method: Tool 5*. <https://www.ahrq.gov/health-literacy/improve/precautions/tool5.html>
- Bean, A. M., Payne, R. E., Parsons, G., Kirk, H. R., Ciro, J., Mosquera-Gómez, R., Hincapié M, S., Ekanayaka, A. S., Tarassenko, L., Rocher, L., & Mahdi, A. (2026). Reliability of LLMs as medical assistants for the general public: A randomized preregistered study. *Nature Medicine*, 32, 609–615. <https://doi.org/10.1038/s41591-025-04074-y>
- Cianciulli, A., Santoro, E., Manente, R., Pacifico, A., Quagliarella, S., Bruno, N., Schettino, V., & Boccia, G. (2025). Artificial intelligence and digital technologies against health misinformation: A scoping review of public health responses. *Healthcare*, 13(20), Article 2623. <https://doi.org/10.3390/healthcare13202623>
- Coalition for Content Provenance and Authenticity. (n.d.). *Content credentials: C2PA technical specification* (Version 2.2). https://spec.c2pa.org/specifications/specification/s/2.2/specs/C2PA_Specification.html
- Combrink, H. M. V. E., & Mkungeka, P. (2025). Misfluencers, the human agents behind AI-driven infodemics. *Journal of Health Communication*, 30(10–12), 349–355. <https://doi.org/10.1080/10810730.2025.2538529>
- Elías-Cabot, E., Romero-Martín, S., Raya-Povedano, J. L., Rodríguez-Ruiz, A., & Álvarez-Benito, M. (2026). AI-based triage and decision support in mammography and digital tomosynthesis for breast cancer screening: A paired, noninferiority trial. *Nature Medicine*, 32, 1296–1305. <https://doi.org/10.1038/s41591-026-04277-x>
- Frederiksen, B., Salganicoff, A., Gomez, I., Diep, K., & Ranji, U. (2026, September 16). *Women's experiences accessing health information online: Trusted sources and contraceptive information*. KFF. <https://www.kff.org/womens-health-policy/womens-experiences-accessing-health-information-online-trusted-sources-and-contraceptive-information/>
- Huang, Q., & Zou, Y. (in press). Understanding older adults' sharing intention toward AI-generated health misinformation: A heuristic-systematic and health belief perspective. *Scientific Reports*. <https://doi.org/10.1038/s41598-026-67991-z>
- Keenan, T., Lampkin, C., Strunk, S., Singer, D., Solway, E., Roberts, S., Box, N., Smith, E., & Kullgren, J. (2024, October). *Health literacy: How well can older adults find, understand, and use health information?* University of Michigan National Poll on Healthy Aging. <https://doi.org/10.7302/24456>
- National Library of Medicine. (2023, October 3). *Evaluating internet health information: A tutorial from the National Library of Medicine: Checklist*. MedlinePlus. <https://medlineplus.gov/webeval/webevalchecklist.html>
- National Library of Medicine. (2025, September 23). *Kidney diseases*. MedlinePlus. <https://medlineplus.gov/kidneydiseases.html>
- National Library of Medicine. (n.d.). *MedlinePlus guidelines for links*. MedlinePlus. <https://medlineplus.gov/about/using/criteria/>
- Novak Jones, D. (2026, May 12). OpenAI faces lawsuit in California court claiming chatbot gave advice that led to fatal overdose. *Reuters*. <https://www.reuters.com/legal/litigation/openai-faces-lawsuit-california-court-claiming-chatbot-gave-advice-that-led-2026-05-12/>
- Office of the Surgeon General. (2021). *A community toolkit for addressing health misinformation*. U.S. Department of Health and Human Services. <https://www.hhs.gov/sites/default/files/health-misinformation-toolkit-english.pdf>
- Office of the Surgeon General. (2021). *Confronting health misinformation: The U.S. Surgeon General's advisory on building a healthy information environment*. U.S. Department of Health and Human Services. <https://www.hhs.gov/sites/default/files/surgeon-general-misinformation-advisory.pdf>
- Omar, M., Sorin, V., Collins, J. D., Reich, D., Freeman, R., Gavin, N., Charney, A., Stump, L., Bragazzi, N. L., Nadkarni, G. N., & Klang, E. (2025). Multi-model assurance analysis showing large language models are highly vulnerable to adversarial hallucination attacks during clinical decision support. *Communications Medicine*, 5, Article 330. <https://doi.org/10.1038/s43856-025-01021-3>
- Omar, M., Sorin, V., Wieler, L. H., Charney, A. W., Kovatch, P., Horowitz, C. R., Korfiatis, P., Glicksberg, B. S., Freeman, R., Nadkarni, G. N., & Klang, E. (2026). Mapping the susceptibility of large language models to medical misinformation across clinical notes and social media: A cross-sectional benchmarking analysis. *The Lancet Digital Health*, 8, Article 100949. <https://doi.org/10.1016/j.landig.2025.100949>
- Presiado, M., Montero, A., Lopes, L., & Hamel, L. (2024, August 15). *KFF health misinformation tracking poll: Artificial intelligence and health information*. KFF. <https://www.kff.org/public-opinion/kff-health-misinformation-tracking-poll-artificial-intelligence-and-health-information/>
- Saeidnia, H. R., Jahani, S., Ghiasi, N., & Keshavarz, H. (2026). Generative AI and health misinformation: Production, propagation, and mitigation—a systematic review. *BMC Public Health*, 26, Article 693. <https://doi.org/10.1186/s12889-025-26148-9>
- Shao, N. (2025). Medical AI, misinformation, and the value of engaging with patients. *CMAJ*, 197, E921–E922. <https://doi.org/10.1503/cmaj.250663>
- Stempel, J. (2026, May 5). Pennsylvania sues Character AI, says chatbot poses as doctors. *Reuters*. <https://www.reuters.com/legal/litigation/pennsylvania-sues-character-ai-says-chatbot-poses-doctors-2026-05-05/>
- Tang, H., Lenzini, G., Greiff, S., Rohles, B., & Sergeeva, A. (2024). “Who knows? Maybe it really works”: Analysing users' perceptions of health misinformation on social media. In *Proceedings of the 2024 ACM Designing Interactive Systems Conference* (pp. 1499–1517). Association for Computing Machinery. <https://doi.org/10.1145/3643834.3661510>
- Tiller, N. B., Marcon, A. R., Zenone, M., Kidd, K. E., Jeukendrup, A. E., Master, Z., & Caulfield, T. (2026). Generative artificial intelligence-driven chatbots and medical misinformation: An accuracy, referencing and readability audit. *BMJ Open*, 16, Article e112695. <https://doi.org/10.1136/bmjopen-2025-112695>
- U.S. Food and Drug Administration. (n.d.). *List of artificial intelligence-enabled medical devices*. <https://www.fda.gov/medical-devices/artificial-intelligence-enabled-medical-devices/list-artificial-intelligence-enabled-medical-devices>
- World Health Organization Regional Office for Europe. (2025). *Responsible AI use can advance risk communication and infodemic management in emergencies, new study shows*. <https://www.who.int/europe/news/item/23-05-2025-responsible-ai-use-can-advance-risk-communication-and-infodemic-management-in-emergencies--new-study-shows>
- World Health Organization. (2024). *WHO releases AI ethics and governance guidance for large multi-modal models*. <https://www.who.int/news/item/18-01-2024-who-releases-ai-ethics-and-governance-guidance-for-large-multi-modal-models>
- Shalu, Verma, N., Dev, K., Bhardwaj, A. B., & Kumar, K. (2025). The Cognitive Cost of AI: How AI Anxiety and Attitudes Influence Decision Fatigue in Daily Technology Use. *Annals of neurosciences*, 09727531251359872. Advance online publication. <https://doi.org/10.1177/09727531251359872>