

Fabricated references are twice as common in medRxiv preprints as in peer-reviewed articles

Dear Editor,

Fabricated references—citations whose claimed titles correspond to no existing publication—have been identified in peer-reviewed biomedical literature at increasing rates [1]. Our recent audit of 2.5 million PubMed Central (PMC) articles identified 4046 fabricated references, with the rate growing more than 12-fold from 2023 to early 2026 [2]. One plausible source of fabrication is large language models (LLMs), which generate plausible but fictitious references; previous studies estimate that 30%–69% of LLM-generated references in biomedical contexts are fabricated [3, 4]. Whether fabricated references also appear in unrefereed preprints, where LLM-assisted drafting is common and reference-checking absent, is unknown, despite the growing role of preprints in clinical evidence synthesis [5, 6].

We adapted the CITADEL reference-verification pipeline [2] for medRxiv preprints posted between January 1, 2023, and March 31, 2026. Because medRxiv references rarely include a PubMed identifier, the primary verification pathway compared each claimed reference title against four scholarly databases covering more than 250 million records (PubMed, Crossref, OpenAlex, and Google Scholar). References not found in any database were flagged as candidate fabrications. After automated screening, 184 candidates underwent independent review by two researchers (P.G. and Z.Z.), each unaware of the other's verdicts. Disagreements were resolved via discussion. This study used publicly posted preprint data and did not require institutional review board approval.

We scanned 42,249 unique medRxiv preprints. The two reviewers agreed on 169 of 184 candidates (91.8% raw agreement; Cohen's kappa, 0.63). Of the 184 candidates reviewed, 157 were confirmed fabricated (precision, 85.3%), yielding 157 fabricated references in 104 preprints (Table 1). The medRxiv rate was 37.2 per 10,000 preprints

(95% confidence interval [CI], 31.6–43.4), more than twice the rate in peer-reviewed PMC articles (16.4 per 10,000; 95% CI, 15.9–16.9) [2], corresponding to a rate ratio of 2.27 (95% CI, 1.94–2.67; $p < 0.001$). Affected preprints were concentrated in medical education (117.0 per 10,000), gastroenterology (86.8 per 10,000), and health systems research (79.5 per 10,000). The annual rate nearly doubled from 2023–2024 to 2025 (approximately 30 per 10,000 in 2023–2024, rising to 57.2 per 10,000 in 2025). Nearly 30% of affected preprints contained two or more fabricated references (a maximum of seven). Of the 104 affected preprints, 35 (34%) had been cited by at least one downstream work, and 12 (12%) had been published in peer-reviewed journals, carrying their fabricated references into the indexed literature. These proportions are conservative because recently posted preprints have had limited time to accrue citations.

Reference errors are common in medical literature [7], but fabricated references differ from routine citation errors because they reference publications that do not exist, potentially creating the appearance of evidence where none exists. We found that validated fabricated references were more than twice as common in medRxiv preprints as in peer-reviewed PMC articles [2], suggesting that peer review and editorial processes may remove some fabricated references before publication, although these safeguards remain incomplete. Our findings complement the unpublished work of Zhao et al. [8], who reported increasing rates of non-existent citations across major scholarly repositories using a different methodology and denominator. Although Zhao et al. quantified repository-wide trends in unmatched references, our study provides a conservative, manually validated estimate of confirmed fabricated references in biomedical preprints, a corpus not separately examined in their analysis. Together, these studies indicate that fabricated citations are an emerging challenge across the scientific publishing ecosystem.

Table 1. Validated fabricated reference rates in medRxiv preprints and PubMed Central, January 2023 to early 2026.

Corpus	Articles scanned	Fabricated references	Affected articles	Rate per 10,000	95% CI
PMC (peer-reviewed)	2,471,758	4046	2810	16.4	15.9–16.9
medRxiv (preprints)	42,249	157	104	37.2	31.6–43.4
Ratio (medRxiv: PMC)	–	–	–	2.27	–

Abbreviations: CI, confidence interval; PMC, PubMed Central.

The demonstrated pathway from preprint to peer-reviewed journal to downstream citation confirms that fabricated references in preprints are not self-contained. Among preprints containing fabricated references, 12 were subsequently published in peer-reviewed journals with the fabricated references intact, and 10 of these published articles accumulated additional citations. This pathway from preprint to journal publication to downstream citation shows how fabricated references can enter and persist within the biomedical evidence ecosystem. Although the absolute rate was low, the volume of preprints used in evidence synthesis means that even uncommon fabrication may create cumulative risk [9].

This concern is increasingly reflected in preprint policy. In May 2026, arXiv announced that submissions containing clear evidence of unchecked LLM-generated content, including hallucinated references, may trigger a 1-year submission ban for authors, followed by additional restrictions on future submissions. Such actions suggest growing recognition that reference integrity is becoming a preprint screening issue, not only a post-publication correction issue [10].

This study has several limitations, including the inability to estimate recall, partial overlap between corpora because some preprints are later indexed in PMC, differing field distributions across corpora, and the absence of causal attribution. Automated reference-integrity checks at preprint submission could help intercept fabricated references before public dissemination.

Author contributions

Concept and design: Maxim Topaz and Laura-Maria Peltonen. *Acquisition, analysis, or interpretation of data:* Maxim Topaz. *Data curation and verification:* Zhihong Zhang and Pallavi Gupta. *Drafting of the manuscript:* Maxim Topaz and Laura-Maria Peltonen. *Critical revision of the manuscript*

for important intellectual content: Maxim Topaz, Zhihong Zhang, Pallavi Gupta, Nir Roguin, and Laura-Maria Peltonen.

Acknowledgments

Two AI tools were used in this work. Claude 3.5 Haiku (Anthropic, San Francisco, CA, USA), accessed April 2026, was used within the CITADEL reference-verification pipeline described in the text to screen flagged references and distinguish genuine fabricated references from formatting discrepancies such as abbreviated titles. ChatGPT (GPT-5.5, OpenAI, San Francisco, CA, USA), accessed May 2026, was used to assist with language editing during manuscript preparation. All AI-generated outputs, including the pipeline's reference classifications and the edited text, were independently reviewed and verified by the authors before inclusion, and the authors take full responsibility for the content of this manuscript.

Conflict of interest statement

The authors declare no conflicts of interest.

Funding information

No funding was received for this work.

Data availability statement

The validated fabricated reference dataset and pipeline code are available upon reasonable request to the corresponding author.

**Maxim Topaz^{1,2,3} , Zhihong Zhang^{1,2} ,
Pallavi Gupta¹ , Nir Roguin⁴ , &
Laura-Maria Peltonen^{5,6,7,8} **

From the ¹*School of Nursing, Columbia University, New York, New York, USA;* ²*Data Science Institute, Columbia University, New York, New York, USA;* ³*VNS Health, New York, New York, USA;* ⁴*Tel Aviv Sourasky Medical Center, Tel Aviv, Israel;* ⁵*Department of Health and Social Management, University of Eastern Finland, Kuopio, Finland;*

⁶Wellbeing Services County of North Savo, Kuopio, Finland; ⁷Wellbeing Services County of Southwest Finland, Turku, Finland; and ⁸Department of Nursing Science, University of Turku, Turku, Finland

References

- 1 Else H, Van Noorden R. The fight against fake-paper factories that churn out sham science. *Nature*. 2021;**591**:516–9.
- 2 Topaz M, Roguin N, Gupta P, Zhang Z, Peltonen L-M. Fabricated citations: an audit across 2.5 million biomedical papers. *Lancet*. 2026;**407**(10541):1779–81.
- 3 Athaluri SA, Manthena SV, Kesapragada VSRKM, Yarlagadda V, Dave T, Duddumpudi RTS. Exploring the boundaries of reality: investigating the phenomenon of artificial intelligence hallucination in scientific writing through ChatGPT references. *Cureus*. 2023;**15**(4):e37432.
- 4 Walters WH, Wilder EI. Fabrication and errors in the bibliographic citations generated by ChatGPT. *Sci Rep*. 2023;**13**(1):14045.
- 5 Fraser N, Brierley L, Dey G, Polka JK, Pálffy M, Nanni F, et al. The evolving role of preprints in the dissemination of COVID-19 research and their impact on the science communication landscape. *PLoS Biol*. 2021;**19**(4):e3000959.
- 6 Flanagan A, Fontanarosa PB, Bauchner H. Preprints involving medical research—do the benefits outweigh the challenges? *J Am Med Assoc*. 2020;**324**(18):1840–43.
- 7 Jergas H, Baethge C. Quotation accuracy in medical journal articles—a systematic review and meta-analysis. *PeerJ*. 2015;**3**:e1364.
- 8 Zhao Z, Wang Y, Stuart T, De Vaan M, Ginsparg P, Yin Y. LLM hallucinations in the wild: large-scale evidence from non-existent citations. *arXiv*. 2026. arXiv:2605.07723. <https://doi.org/10.48550/arXiv.2605.07723>
- 9 Tang G, Cai H. Citation contamination by paper mill articles in systematic reviews of the life sciences. *JAMA Netw Open*. 2025;**8**:e2515160.
- 10 Singh Chawla D. Researchers who use hallucinated references to face arXiv ban. *Nature*. 2026;**653**:988–9.

Correspondence: Maxim Topaz, School of Nursing, Columbia University, 560 West 168th Street, New York, NY 10032, USA.
Email: mt3315@cumc.columbia.edu