

Inteligencia artificial para artículos neuroquirúrgicos

La irrupción de modelos de lenguaje basados en inteligencia artificial (IA) como ChatGPT ha comenzado a transformar de manera sustancial la escritura académica en neurocirugía. Diversos estudios recientes sugieren que estos sistemas pueden mejorar la legibilidad de los textos científicos, facilitar la redacción para autores no nativos en inglés, automatizar partes del proceso editorial y actuar como herramientas pedagógicas para residentes. Sin embargo, esta aparente revolución plantea interrogantes críticos sobre la profundidad analítica, la autoría, la integridad científica y la ética editorial.

Fauziah et al. (2025) demostraron que los artículos generados por IA obtuvieron puntuaciones de legibilidad significativamente superiores (Lix: 35 vs. 26; Flesch-Kincaid: 10 vs. 8) respecto a los textos humanos, aunque con menor profundidad en el análisis. Sorprendentemente, evaluadores expertos solo identificaron correctamente la autoría de los textos en el 61 % de los casos, y la preferencia entre artículos generados por IA (47 %) y humanos (44 %) fue prácticamente equitativa ¹⁾. Estos hallazgos apuntan a un creciente solapamiento entre el “formato” del texto y su “fondo”.

En una línea similar, Akgun et al. (2025) compararon cinco artículos generados por ChatGPT con cinco escritos por humanos y encontraron calidad global comparable entre ambos grupos. De nuevo, las puntuaciones de legibilidad favorecieron a la IA, y la preferencia de los expertos estuvo dividida (47 % IA, 44 % humano) ²⁾. No obstante, ambos estudios carecen de una auditoría sistemática de la veracidad clínica de los contenidos, lo cual es especialmente crítico en una disciplina como la neurocirugía.

Schneider et al. (2025) adoptaron un enfoque más técnico para evaluar la prevalencia de contenido generado por IA en 100 manuscritos de revistas neuroquirúrgicas. Utilizando una combinación de un clasificador RoBERTa y un umbral de perplexity <100, identificaron que el 20 % de los artículos contenían secciones probablemente redactadas con IA, particularmente en resúmenes y métodos. Al excluir los resúmenes, el efecto dejó de ser estadísticamente significativo, lo cual sugiere una posible sobreestimación en secciones altamente estructuradas ³⁾. A destacar, los autores recomiendan abandonar la obsesión por la detección y centrar los esfuerzos en políticas de divulgación y transparencia.

Nabata et al. (2025), por su parte, exploraron la capacidad de profesionales quirúrgicos para distinguir entre resúmenes científicos generados por ChatGPT y originales. Solo el 40 % de los participantes identificó correctamente el resumen humano, mientras que un 63 % prefirió el generado por IA ($p = 0.0001$). El hecho de que incluso cirujanos experimentados no puedan diferenciar el origen de un texto apunta a la creciente sofisticación de estos modelos, pero también a un posible desplazamiento de la percepción de calidad hacia la forma sobre el contenido ⁴⁾.

En conjunto, estos estudios delinean un nuevo escenario donde la IA puede servir como **herramienta de apoyo poderosa pero no como sustituto del criterio clínico, la experiencia investigadora y la capacidad de análisis crítico**. La legibilidad no debe confundirse con comprensión, y la coherencia formal no garantiza rigor metodológico ni validez científica.

El futuro de la escritura académica en neurocirugía debería contemplar la IA como un **coautor menor, transparente y regulado**, al estilo del software estadístico o los gestores bibliográficos. Además, urge diseñar guías editoriales que regulen la atribución de autoría, la responsabilidad sobre

el contenido y la validación de datos clínicos cuando se utilicen estas herramientas. La formación en IA aplicada a la redacción científica debe incorporarse a los programas de residencia con una perspectiva ética y crítica.

En suma, la IA puede mejorar la eficiencia y accesibilidad del proceso editorial, pero su integración debe realizarse con **responsabilidad epistemológica y ética profesional**, evitando que la forma eclipse el fondo en la producción científica.

Referencias

1. ⁵⁾ Fauziah RR, Puspita AMI, Yuliana I, Ummah FS, Mufarochah S, Ramadhani E. Artificial intelligence in academic writing: Enhancing or replacing human expertise? *J Clin Neurosci*. 2025 Mar 19;111193. doi:10.1016/j.jocn.2025.111193.
2. ⁶⁾ Schneider DM, Mishra A, Gluski J, Shah H, Ward M, Brown ED, Sciubba DM, Lo SL. Prevalence of Artificial Intelligence-Generated Text in Neurosurgical Publications: Implications for Academic Integrity and Ethical Authorship. *Cureus*. 2025 Feb 16;17(2):e79086. doi:10.7759/cureus.79086.
3. ⁷⁾ Akgun MY, Savasci M, Gunerbuyuk C, Gunara SO, Oktenoglu T, Ozer AF, Ates O. Battle of the authors: Comparing neurosurgery articles written by humans and AI. *J Clin Neurosci*. 2025 Feb 25;135:111152. doi:10.1016/j.jocn.2025.111152.
4. ⁸⁾ Nabata KJ, AlShehri Y, Mashat A, Wiseman SM. Evaluating human ability to distinguish between ChatGPT-generated and original scientific abstracts. *Updates Surg*. 2025 Jan 24. doi:10.1007/s13304-025-02106-3.

1) ⁵⁾

,

1

2) ⁷⁾

,

3

3) ⁶⁾

,

2

4) ⁸⁾

,

4

From:

<https://neurocirugiacontemporanea.es/wiki/> - Neurocirugía Contemporánea ISSN 1988-2661

Permanent link:

https://neurocirugiacontemporanea.es/wiki/doku.php?id=inteligencia_artificial_para_articulos_neuroquirurgicos

Last update: 2025/05/03 23:57

