

# Inteligencia artificial en salud mental: evaluación temprana del riesgo suicida con responsabilidad ética

*Artificial intelligence in mental health: early assessment of suicide risk with ethical responsibility*

**Jorge Hernández Navas<sup>1</sup>, Luis Dulcey Sarmiento<sup>2</sup>, Jaime Gómez Ayala<sup>3</sup>, Juan Therán León<sup>4</sup>, Javier Alonso Vesga<sup>5</sup>, Oscar Daniel López<sup>6</sup>**

<https://doi.org/10.53680/vertex.v36i170.949>

## Resumen

El suicidio constituye una de las principales causas de muerte prevenible a nivel global, afectando especialmente a personas jóvenes entre los 15 y 29 años. Esta problemática de salud pública tiene un impacto devastador en individuos, familias y comunidades, y plantea un desafío clínico persistente para los sistemas de salud. A pesar de décadas de investigación en prevención, los profesionales continúan enfrentando dificultades considerables para identificar de manera temprana a los individuos en riesgo inminente. En este escenario, la inteligencia artificial (IA) ha surgido como una herramienta innovadora con el potencial de transformar la evaluación y detección del riesgo suicida mediante el análisis de datos complejos y multimodales.

**Palabras clave:** suicidio, inteligencia artificial, chatbots

## Abstract

*Suicide is one of the leading causes of preventable death globally, particularly affecting young people between the ages of 15 and 29. This public health problem has a devastating impact on individuals, families, and communities, and poses a persistent clinical challenge for health systems. Despite decades of research in prevention, professionals continue to face considerable difficulties in identifying individuals at imminent risk early on. In this context, artificial intelligence (AI) has emerged as an innovative tool with the potential to transform suicide risk assessment and detection through the analysis of complex and multimodal data.*

**Keywords:** suicide, artificial intelligence, chatbots

RECIBIDO 9/5/2025 - ACEPTADO 6/8/2025

<sup>1</sup>Médico investigador. Universidad de Santander, Bucaramanga, Colombia. <https://orcid.org/0009-0001-5758-5965>

<sup>2</sup>Médico especialista en medicina interna. Universidad de los Andes, Mérida, Venezuela. <https://orcid.org/0000-0001-9306-0413>

<sup>3</sup>Profesor titular. Facultad de medicina. Universidad Autónoma de Bucaramanga, Colombia. <https://orcid.org/0000-0002-1103-9598>

<sup>4</sup>Médico residente de medicina familiar. Universidad de Santander, Bucaramanga, Colombia. <https://orcid.org/0000-0002-4742-0403>

<sup>5</sup>Médico general. Universidad Industrial de Santander, Bucaramanga, Colombia. <https://orcid.org/0000-0001-8956-2912>

<sup>6</sup>Médico general. Universidad de Santander, Bucaramanga, Colombia. <https://orcid.org/0000-0002-7280-3086>

## Autor correspondiente:

Jorge Andrés Hernández Navas

[Jorgeandreshernandez2017@gmail.com](mailto:Jorgeandreshernandez2017@gmail.com)



## Introducción

El suicidio constituye una de las principales causas de muerte prevenible a nivel global, afectando especialmente a personas jóvenes entre los 15 y 29 años. Esta problemática de salud pública tiene un impacto devastador en individuos, familias y comunidades, y plantea un desafío clínico persistente para los sistemas de salud. A pesar de décadas de investigación en prevención, los profesionales continúan enfrentando dificultades considerables para identificar de manera temprana a los individuos en riesgo inminente. En este escenario, la Inteligencia Artificial (IA) ha surgido como una herramienta innovadora con el potencial de transformar la evaluación y detección del riesgo suicida mediante el análisis de datos complejos y multimodales (Dehbozorgi et al., 2025).

## Detección de riesgo suicida con IA

El uso de algoritmos de aprendizaje automático permite procesar grandes volúmenes de información estructurada y no estructurada, tales como historias clínicas electrónicas, notas de evolución, patrones de uso de servicios de urgencias, registros de antecedentes psiquiátricos y, más recientemente, datos de comportamiento digital en redes sociales. Estos modelos han mostrado una capacidad destacada para identificar patrones sutiles y no evidentes para el juicio clínico humano, lo cual podría facilitar una detección más oportuna de pacientes con riesgo elevado. Por ejemplo, el análisis semántico de lenguaje natural – tanto oral como escrito – ha demostrado su utilidad para detectar expresiones indirectas de desesperanza, anhedonia, ideación suicida encubierta o aislamiento social, incluso en ausencia de una declaración explícita de intenciones autolíticas (Dehbozorgi et al., 2025; Cruz-Gonzalez et al., 2025).

## Consideraciones éticas

No obstante, este progreso técnico requiere una reflexión ética y crítica profunda. La aplicación de sistemas de IA en salud mental plantea una serie de desafíos fundamentales relacionados con la equidad, la privacidad, el consentimiento y la autonomía de los pacientes. Uno de los principales riesgos consiste en la posibilidad de que los algoritmos reproduzcan y perpetúen sesgos preexistentes en los datos de entrenamiento, especialmente cuando estos reflejan desigualdades estructurales vinculadas a factores sociales, económicos, raciales o de género. Tal situación podría conducir a la sobrerepresentación o estigmatización

de determinados grupos, generando consecuencias clínicas y sociales adversas (Dakanalis et al., 2024; Cruz-Gonzalez et al., 2025).

A ello se suma la preocupación por la opacidad de muchos modelos algorítmicos, conocidos como “cajas negras”, cuya lógica de funcionamiento no es fácilmente comprensible ni auditabile por profesionales clínicos. Esta falta de explicabilidad dificulta la validación de las decisiones asistidas por IA, compromete su confiabilidad ética y pone en entredicho la posibilidad de una supervisión médica adecuada. Además, el uso de datos sensibles de salud mental plantea un desafío adicional en cuanto a la protección de la privacidad y la confidencialidad del paciente, especialmente cuando se exploran fuentes de información no tradicionales como el contenido de redes sociales o históricas de navegación (Dakanalis et al., 2024).

La IA podría revolucionar la prevención del intento suicida al integrar múltiples factores que trascienden la anamnesis clínica tradicional. En Latinoamérica, investigaciones como las del grupo de investigación GIS en Bogotá, Colombia, han señalado que las redes sociales constituyen un espacio donde las personas expresan de mayor forma sus emociones, pensamientos y preocupaciones. Abarcando y explicando como los matices en Colombia van aumentando los riesgos dependiendo de su parte socioeconómica y de los grupos etarios de riesgo como lo son los adolescentes. Siendo así una herramienta para la detección de señales tempranas no visibles o que no son evaluables en una consulta como cambios en el lenguaje, menciones de términos asociados al suicidio, o sentimientos de desesperanza y aislamiento en sus redes sociales (Carmelo Alejo et al., 2025; García-López et al., 2023).

## Conclusiones

En este contexto, resulta imprescindible que la integración de la IA en la evaluación del riesgo suicida se rija por un marco normativo sólido, basado en principios de bioética y derechos humanos. Dicho marco debe contemplar mecanismos rigurosos de consentimiento informado, asegurar la supervisión profesional permanente, establecer protocolos para la intervención ante alertas generadas por algoritmos, y garantizar la transparencia y la auditabilidad de las herramientas utilizadas. Del mismo modo, se debe promover la inclusión activa de profesionales de salud mental en el diseño, validación y monitoreo de estos sistemas, así como la participación de comités de ética, juristas, y representantes de pacientes o usuarios del sistema de

salud, a fin de asegurar la pertinencia cultural y contextual de las soluciones tecnológicas desarrolladas (Sauerbrei et al., 2023; Dakanalis et al., 2024).

La IA no puede ni debe sustituir la relación terapéutica basada en la empatía, la escucha activa y el juicio clínico informado. Su utilidad radica en actuar como una herramienta de apoyo que amplía las capacidades diagnósticas y preventivas del sistema de salud, permitiendo una respuesta más ágil y personalizada ante situaciones de riesgo (Ghadiri et al., 2024). Se ha evidenciado que estas herramientas predictivas con IA han logrado avances considerables. Se ha reportado que estas herramientas predictivas alcanzan tasas de detección superiores al 90 %, como en el diagnóstico de trastorno depresivo y Enfermedad de Alzheimer (Ocando Parra, 2025).

La IA, a través de recursos como los chatbots terapéuticos, ofrece ventajas adicionales, al favorecer el anonimato y facilitar una comunicación más abierta y permanente con los pacientes. Además, beneficiando a un entorno donde se tiene una gran demanda de servicios psiquiátricos sin disponibilidad de especialistas para cubrir todas las regiones.

En conclusión, la IA representa una oportunidad prometedora para reforzar los esfuerzos en la prevención del suicidio, al facilitar la identificación temprana de individuos en riesgo y posibilitar intervenciones más oportunas y eficaces. Este enfoque permite pasar de una atención única en un consultorio hacia un modelo de seguimiento continuo, que ofrece a los pacientes mayores oportunidades de apertura y la posibilidad de identificar patrones de riesgo menos evidentes. Sin embargo, este potencial solo podrá concretarse si su aplicación se enmarca dentro de una ética del cuidado, informada por la evidencia científica, sensible a las realidades sociales y comprometida con la defensa de la persona. De lo contrario, existe el riesgo de que la tecnología, en lugar de aliviar el sufrimiento, se convierta en una nueva forma de exclusión, vigilancia y violencia estructural (Ocando Parra, 2025). En ese escenario, el paciente podría reducirse a un simple conjunto de datos, despojado de su dimensión huma-

na. Por ello, cualquier avance técnico en este campo debe ir acompañado de una profunda reflexión ética, un diálogo interdisciplinario y un compromiso firme con los valores fundamentales de la salud mental comunitaria (Ghadiri et al., 2024; Alowais et al., 2023).

**Conflictos de interés:** los autores no declaran conflictos de interés.

## Referencias bibliográficas

- Dehbozorgi, R., Zangeneh, S., Khooshab, E., Nia, D. H., Hanif, H. R., Samian, P., Yousefi, M., Hashemi, F. H., Vakili, M., Jamalioghadam, N., & Lohrasebi, F. (2025). The application of artificial intelligence in the field of mental health: a systematic review. *BMC psychiatry*, 25(1), 132. <https://doi.org/10.1186/s12888-025-06483-2>
- Cruz-Gonzalez, P., He, A. W., Lam, E. P., Ng, I. M. C., Li, M. W., Hou, R., Chan, J. N., Sahnii, Y., Vinas Guasch, N., Miller, T., Lau, B. W., & Sánchez Vidaña, D. I. (2025). Artificial intelligence in mental health care: a systematic review of diagnosis, monitoring, and intervention applications. *Psychological medicine*, 55, e18. <https://doi.org/10.1017/S0033291724003295>
- Dakanalis, A., Wiederhold, B. K., & Riva, G. (2024). Artificial Intelligence: A Game-Changer for Mental Health Care. *Cyberpsychology, behavior and social networking*, 27(2), 100–104. <https://doi.org/10.1089/cyber.2023.0723>
- Camelo Alejo, L., García Alfonso, D., Montaño Gutiérrez, J., Romero Cuellar, L. Aplicación de inteligencia artificial para la prevención de intentos de suicidio en Bogotá. [Internet]. Universidad EAN; 2025 [citado: 2025, agosto] 47 páginas.
- García-López, A., Girón-Luque, F., Rosselli, D. La integración de la inteligencia artificial en la atención médica: desafíos éticos y de implementación. Revistasjaverianaeduo [Internet]. 2023 May 24 [cited 2025 Aug 16]; Available from: [https://revistas.javeriana.edu.co/files-articulos/UMED/64-3\(2023\)/6572567006/index.html](https://revistas.javeriana.edu.co/files-articulos/UMED/64-3(2023)/6572567006/index.html)
- Sauerbrei, A., Kerasidou, A., Lucivero, F., & Hallowell, N. (2023). The impact of artificial intelligence on the person-centred, doctor-patient relationship: some problems and solutions. *BMC medical informatics and decision making*, 23(1), 73. <https://doi.org/10.1186/s12911-023-02162-y>
- Ghadiri, P., Yaffe, M. J., Adams, A. M., & Abbasgholizadeh-Rahimi, S. (2024). Primary care physicians' perceptions of artificial intelligence systems in the care of adolescents' mental health. *BMC primary care*, 25(1), 215. <https://doi.org/10.1186/s12875-024-02417-1>
- Ocando Parra, L. C. (2025). Inteligencia artificial en psiquiatría: innovaciones, desafíos y futuro del diagnóstico y tratamiento. Revisión bibliográfica. *Revista Latinoamericana de Hipertensión*, 20(1):85–94.
- Alowais, S. A., Alghamdi, S. S., Alsuhebany, N., Alqahtani, T., Alshaya, A. I., Almohareb, S. N., Aldairem, A., Alrashed, M., Bin Saleh, K., Badreldin, H. A., Al Yami, M. S., Al Harbi, S., & Albekairy, A. M. (2023). Revolutionizing healthcare: the role of artificial intelligence in clinical practice. *BMC medical education*, 23(1), 689. <https://doi.org/10.1186/s12909-023-04698-z>

volver al índice