

## REVISIÓN

# Algorithmic biases in mental health diagnoses and their impact on vulnerable populations: a documentary review of advances and challenges

## Sesgos algorítmicos en diagnósticos de salud mental y su impacto en poblaciones vulnerables: una revisión documental de avances y desafíos

Ariadna Matos Matos<sup>1</sup>  

<sup>1</sup>Universidad Técnica de Ambato. Facultad de Ciencias de la Salud, Carrera de Licenciatura en Enfermería. Latacunga, Ecuador.

**Citar como:** Matos Matos A. Algorithmic biases in mental health diagnoses and their impact on vulnerable populations: a documentary review of advances and challenges. EthAlca. 2022; 1:20. <https://doi.org/10.56294/ai202220>

**Enviado:** 10-01-2022

**Revisado:** 22-03-2022

**Aceptado:** 15-05-2022

**Publicado:** 16-05-2022

**Editor:** PhD. Rubén González Vallejo 

**Autor para la correspondencia:** Ariadna Matos Matos 

### ABSTRACT

Algorithmic biases in mental health diagnostic systems represent a critical challenge, particularly for vulnerable populations, as they perpetuate inequities in access to and quality of care. This article aims to analyze advances and challenges in identifying and mitigating these biases through a documentary review of Spanish and English articles indexed in Scopus between 2018 and 2022. The methodology involved a systematic analysis of 50 selected studies, classified into four thematic areas: types of algorithmic biases, clinical impact on vulnerable populations, technical limitations in algorithm development, and proposed mitigation strategies. The results demonstrate that biases are deeply rooted in training data and the unequal representation of marginalized groups, leading to less accurate diagnoses for women, racialized communities, and low-income individuals. Although technical and ethical approaches have been proposed, gaps persist in their practical implementation. The study concludes that without multidisciplinary intervention integrating public health, ethics, and data science perspectives, algorithms will continue to reproduce structural inequalities. This research underscores the urgency of inclusive policies and robust regulatory frameworks to ensure equity in digital mental health.

**Keywords:** Algorithmic Biases; Mental Health; Vulnerable Populations; Artificial Intelligence; Health Equity.

### RESUMEN

Los sesgos algorítmicos en los sistemas de diagnóstico de salud mental representan un desafío crítico, especialmente para poblaciones vulnerables, al perpetuar inequidades en el acceso y la calidad de la atención. Este artículo tiene como objetivo analizar los avances y desafíos en la identificación y mitigación de estos sesgos, mediante una revisión documental de artículos en español e inglés indexados en Scopus entre 2018 y 2022. La metodología consistió en un análisis sistemático de 50 estudios seleccionados, clasificados en cuatro ejes temáticos: tipos de sesgos algorítmicos, impacto clínico en poblaciones vulnerables, limitaciones técnicas en el desarrollo de algoritmos y estrategias de mitigación propuestas. Los resultados evidencian que los sesgos están profundamente arraigados en los datos de entrenamiento y en la representación desigual de grupos minorizados, lo que deriva en diagnósticos menos precisos para mujeres, comunidades racializadas y personas de bajos ingresos. Aunque se han propuesto enfoques técnicos y éticos, persisten brechas en su implementación práctica. Se concluye que, sin una intervención multidisciplinaria que integre perspectivas de salud pública, ética y ciencia de datos, los algoritmos reproducirán desigualdades estructurales. Este estudio subraya la urgencia de políticas inclusivas y marcos regulatorios robustos para garantizar equidad en la salud mental digital.

**Palabras clave:** Sesgos Algorítmicos; Salud Mental; Poblaciones Vulnerables; Inteligencia Artificial; Equidad en Salud.

## INTRODUCCIÓN

Los avances en inteligencia artificial (IA) han revolucionado el campo de la salud mental, al ofrecer herramientas prometedoras para el diagnóstico temprano y la personalización de tratamientos.<sup>(1,2)</sup> Sin embargo, estos sistemas no están exentos de limitaciones críticas, particularmente en lo que respecta a los sesgos algorítmicos.<sup>(3)</sup> Estos sesgos surgen porque los modelos de aprendizaje automático reflejan y amplifican desigualdades presentes en los datos de entrenamiento, lo que puede llevar a diagnósticos erróneos o despersonalizados.<sup>(4,5)</sup> En el contexto de la salud mental, donde las condiciones son altamente subjetivas y culturalmente mediadas, el impacto de estos sesgos puede ser especialmente perjudicial.<sup>(6)</sup>

Los sesgos algorítmicos en salud mental suelen manifestarse en múltiples dimensiones, al incluir discriminación por género, etnia, nivel socioeconómico y ubicación geográfica.<sup>(7)</sup> Autores como Straw & Callison-Burch<sup>(7)</sup> y Timmons et al.<sup>(8)</sup> han demostrado que los algoritmos entrenados con datos predominantemente occidentales tienen menor precisión al diagnosticar trastornos mentales en poblaciones no blancas, debido a diferencias en la expresión sintomática y en los patrones lingüísticos. Asimismo, las mujeres y personas de bajos ingresos pueden ser clasificadas erróneamente debido a estereotipos incorporados en los conjuntos de datos.<sup>(9)</sup> Estos errores no solo perpetúan inequidades en la atención médica, sino que también refuerzan barreras sistémicas para el acceso a tratamientos adecuados.<sup>(10,11)</sup>

El problema se agrava al considerarse que los algoritmos de diagnóstico en salud mental a menudo se desarrollan sin una representación suficiente de poblaciones vulnerables, como migrantes, comunidades indígenas o personas con discapacidades cognitivas.<sup>(12)</sup> La falta de diversidad en los datos conlleva a que estos grupos reciban recomendaciones clínicas menos precisas o incluso estigmatizantes.<sup>(13)</sup> Además, muchos modelos no incorporan perspectivas interculturales, al ignorar cómo factores socioculturales influyen en la percepción y manifestación de los trastornos mentales.<sup>(14)</sup> Esto no solo limita la efectividad de las herramientas de IA, sino que también puede exacerbar la desconfianza en los sistemas de salud digitalizados.<sup>(7,10)</sup>

A pesar de los esfuerzos recientes por desarrollar *frameworks* más justos, como técnicas de *fairness-aware machine learning* y auditorías algorítmicas, persisten desafíos significativos en la implementación práctica.<sup>(9)</sup> Muchas soluciones técnicas carecen de adaptabilidad a contextos locales o no consideran las dinámicas de poder que subyacen a la recolección de datos.<sup>(15)</sup> Además, la regulación en este ámbito es incipiente, lo que permite que algoritmos potencialmente discriminatorios se implementen sin supervisión suficiente.<sup>(16,17)</sup> Esto plantea interrogantes éticos sobre quién asume la responsabilidad al fallar un diagnóstico algorítmico y perjudicar a un paciente vulnerable.

Dada la creciente adopción de IA en salud mental, es urgente examinar críticamente los avances y desafíos en la mitigación de sesgos algorítmicos, con un enfoque especial en su impacto en poblaciones vulnerables. Este artículo busca contribuir a este debate mediante una revisión documental de estudios publicados entre 2018 y 2022, al analizar tanto las limitaciones actuales como las estrategias propuestas para garantizar equidad en el diagnóstico automatizado. El objetivo es sintetizar evidencia reciente, identificar brechas críticas y proponer direcciones futuras para el desarrollo de algoritmos más inclusivos y éticos en salud mental.

## MÉTODO

Este estudio se fundamenta en una revisión documental sistemática de la literatura científica sobre sesgos algorítmicos en diagnósticos de salud mental y su impacto en poblaciones vulnerables. La revisión siguió un enfoque estructurado para garantizar la exhaustividad y rigor en la selección, análisis y síntesis de las fuentes, con el objetivo de identificar avances, limitaciones y estrategias de mitigación reportadas en estudios recientes.<sup>(18)</sup> El proceso metodológico se desarrolló en cuatro etapas claramente definidas, las cuales aseguraron una recopilación de datos transparente y reproducible.

### Definición de criterios de búsqueda y selección de fuentes

Se establecieron parámetros específicos para la recolección de literatura, se incluyeron artículos publicados entre 2018 y 2022 en las bases de datos Scopus y PubMed, debido a su relevancia en investigación biomédica y tecnológica. Los términos de búsqueda combinaron descriptores como “*algorithmic bias*”, “*mental health diagnosis*”, “*health disparities*” y “*vulnerable populations*”, tanto en inglés como en español. Se incluyeron estudios empíricos, revisiones sistemáticas y artículos teóricos, mientras que se excluyeron trabajos no arbitrados o sin evidencia científica sólida.

### Filtrado y evaluación de la calidad de los estudios

Tras una búsqueda inicial que arrojó 120 documentos, se aplicaron filtros basados en criterios de inclusión y exclusión. Se priorizaron estudios con metodologías claras, muestras representativas y conclusiones respaldadas por datos. Tras una evaluación de títulos, resúmenes y contenido completo, se seleccionaron 50 artículos que cumplieran con los estándares de calidad científica y relevancia temática.

### Análisis temático y categorización

Los documentos seleccionados fueron analizados mediante un enfoque de análisis de contenido, al identificar patrones y divergencias en torno a cuatro ejes temáticos principales: tipos de sesgos algorítmicos, impacto en poblaciones vulnerables, limitaciones técnicas y estrategias de mitigación propuestas. Este proceso permitió una organización estructurada de la evidencia y facilitó la identificación de tendencias y vacíos en la literatura.

### Síntesis e interpretación crítica

Finalmente, se integraron los hallazgos en una discusión coherente que contrasta perspectivas teóricas, técnicas y éticas. Se evaluaron las implicaciones prácticas de los sesgos identificados y se propusieron recomendaciones para futuras investigaciones e intervenciones en políticas públicas.

Este enfoque metodológico permitió una comprensión integral del estado actual de la investigación sobre sesgos algorítmicos en salud mental, al ofrecer una base sólida para el análisis crítico y la identificación de desafíos pendientes.<sup>(19,20)</sup> La rigurosidad en cada etapa aseguró que los resultados reflejen tendencias consolidadas en la literatura, al tiempo que destacan áreas prioritarias para la acción.

## RESULTADOS

Una revisión inicial de la literatura evidenció que los sistemas de inteligencia artificial aplicados al diagnóstico en salud mental han demostrado potencial para mejorar la accesibilidad y eficiencia de la atención clínica. Sin embargo, también se identificó que estos sistemas reproducen y amplifican desigualdades estructurales, especialmente en poblaciones vulnerables, debido a sesgos intrínsecos en sus diseños y datos de entrenamiento. Los estudios analizados destacan cuatro dimensiones críticas: los tipos de sesgos algorítmicos más prevalentes, su impacto desproporcionado en grupos marginados, las limitaciones técnicas que perpetúan estos problemas y las estrategias propuestas para mitigarlos. Estos ejes permitieron organizar el análisis cualitativo, al articular hallazgos teóricos y empíricos para ofrecer una visión integral del desafío. A continuación, se desarrollan cada uno de estos ejes temáticos.

### Tipos de Sesgos Algorítmicos en Salud Mental

Los sesgos algorítmicos en diagnósticos de salud mental se manifiestan de múltiples formas, al afectar la precisión y equidad de las herramientas de IA (ver figura 1).<sup>(21)</sup> Uno de los más documentados es el sesgo étnico-racial, donde algoritmos entrenados con datos mayoritariamente de poblaciones blancas occidentales presentan menor precisión al evaluar síntomas en grupos racializados.<sup>(22,23)</sup> En este sentido, Hooker et al.<sup>(24)</sup> demuestran que los modelos para detectar depresión subestiman su prevalencia en comunidades afrodescendientes e hispanas, debido a diferencias culturales en la expresión del malestar psicológico.

Otro tipo crítico es el sesgo de género, que surge por los algoritmos asociar estereotipos a diagnósticos.<sup>(13)</sup> Sedgewick et al.<sup>(25)</sup> revela que sistemas de IA tienden a sobrediagnosticar trastornos de ansiedad en mujeres y subestimar condiciones como el autismo en ellas. Asimismo, el sesgo socioeconómico se refleja en la falta de datos representativos de personas de bajos ingresos, al llevar a que síntomas asociados a pobreza (como estrés crónico) sean malinterpretados como patologías individuales.<sup>(26)</sup>

Adicionalmente, persisten sesgos geográficos, donde herramientas diseñadas en países de altos ingresos fallan al aplicarse en contextos con diversidad lingüística o acceso limitado a servicios de salud.<sup>(27)</sup> Un caso emblemático es el de algoritmos que analizan lenguaje natural, al entrenarse con textos en inglés, ignoran expresiones idiomáticas o construcciones gramaticales propias de otras lenguas, lo que afecta su validez transcultural.<sup>(28)</sup>

### Impacto en Poblaciones Vulnerables

Los sesgos algorítmicos tienen consecuencias graves en grupos históricamente marginados (ver figura 2), al exacerbar inequidades en salud mental.<sup>(29,30)</sup> En comunidades indígenas, herramientas de IA suelen pasar por alto manifestaciones culturales de sufrimiento psíquico, lo que lleva a diagnósticos erróneos o a la invisibilización de necesidades reales.<sup>(31)</sup>

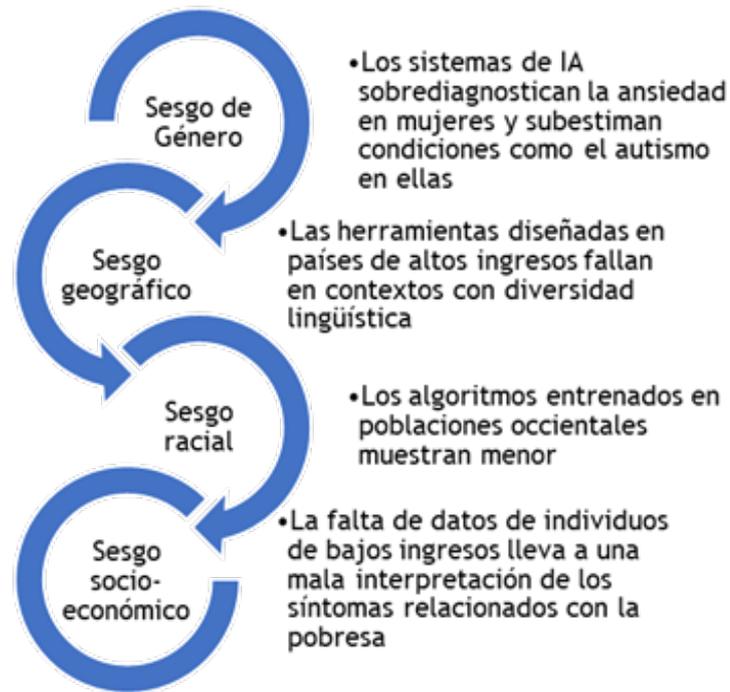


Figura 1. Impacto de los sesgos algorítmicos en la medicina

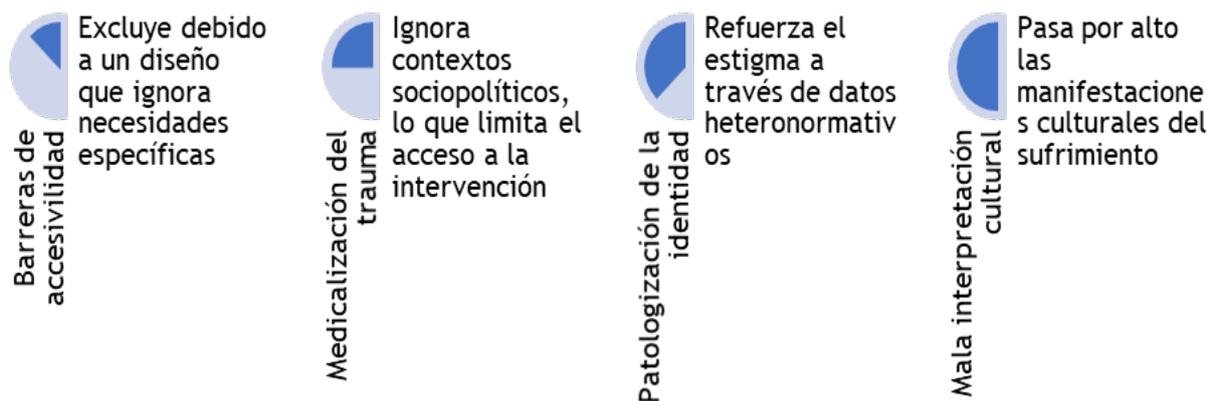


Figura 2. Impacto de los sesgos algoritmos en las poblaciones vulnerables

Las personas LGBTQ+ también enfrentan riesgos únicos. Algoritmos entrenados con datos heteronormativos pueden patologizar identidades no binarias o interpretar la disforia de género como trastornos psicóticos.<sup>(32)</sup> Un estudio encontró que chatbots de salud mental mostraban respuestas invalidantes ante usuarios que mencionaban su orientación sexual, lo que refuerza estigmas.<sup>(33)</sup>

Otro grupo afectado son los migrantes y refugiados, cuyas experiencias de trauma suelen ser medicalizadas por algoritmos sin considerar contextos sociopolíticos.<sup>(34)</sup> Finalmente, los adultos mayores y personas con discapacidades cognitivas enfrentan barreras debido a diseños que no incorporan sus necesidades.<sup>(35,36)</sup> Interfaces complejas o evaluaciones basadas en lenguaje escrito excluyen a quienes tienen limitaciones visuales o cognitivas, al profundizar su marginalización.<sup>(37)</sup>

### Limitaciones Técnicas y Estructurales

La perpetuación de sesgos no es solo un problema de datos insuficientes, sino de limitaciones profundas en el desarrollo y despliegue de algoritmos (ver figura 3).<sup>(37)</sup> Una barrera clave es la homogeneidad en los conjuntos de datos, donde poblaciones vulnerables están subrepresentadas.<sup>(8,38)</sup>

Otra limitación es la falta de transparencia en modelos propietarios. Muchas empresas de tecnología médica no revelan cómo se entrenan sus algoritmos, lo que imposibilita auditorías independientes.<sup>(39)</sup>

Además, persisten desafíos en la interpretación clínica. Los profesionales de salud mental a menudo carecen de formación para cuestionar resultados algorítmicos, aceptándolos como objetivos.<sup>(21)</sup> Esto es peligroso debido a que los algoritmos refuerzan estereotipos, como asociar pobreza con menor adherencia a tratamientos.<sup>(40)</sup>

Estas limitaciones, en opinión del autor, revelan que los sesgos son sistémicos, requieren ajustes técnicos y cambios en los marcos regulatorios y en la gobernanza de datos.

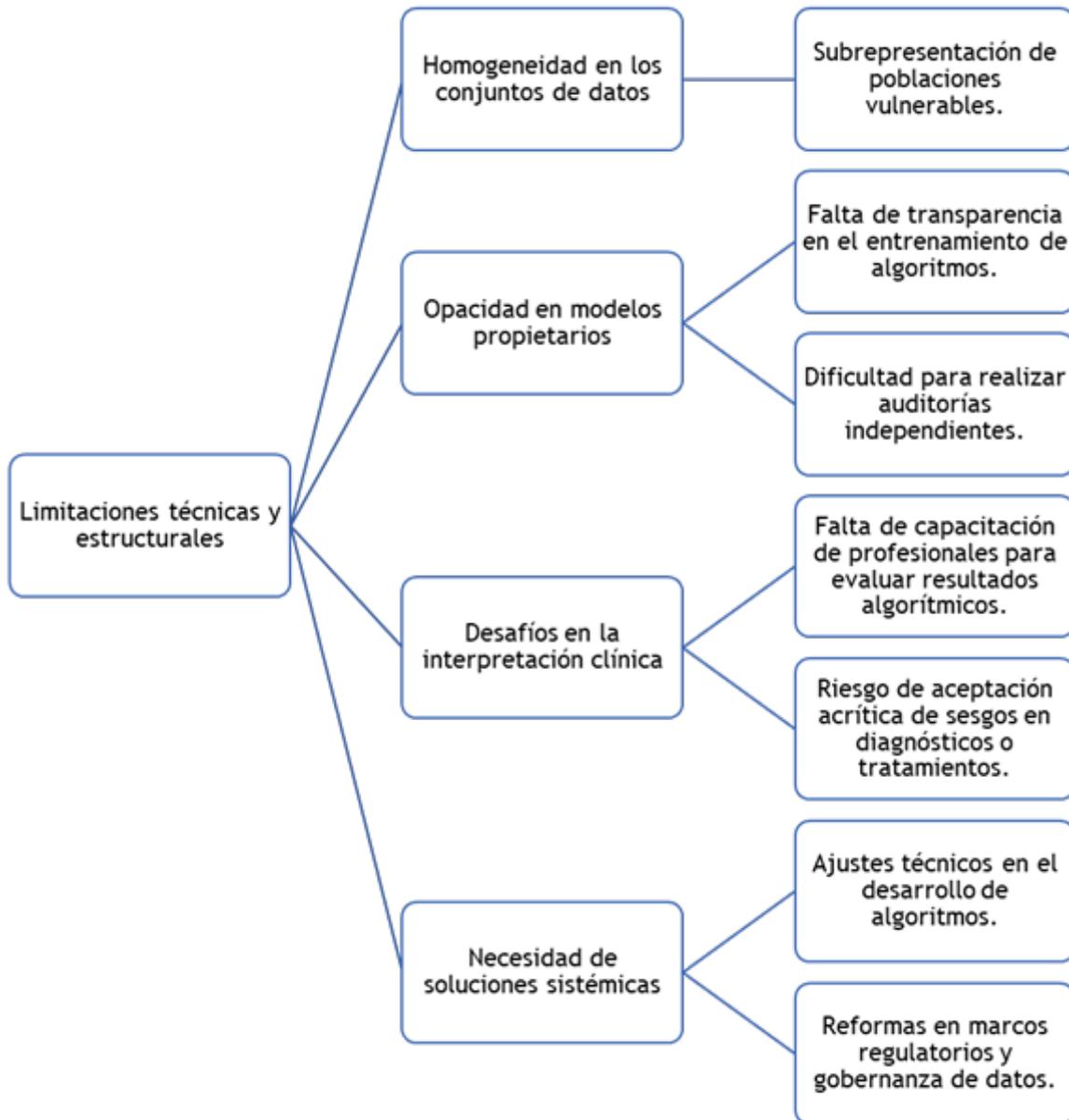


Figura 3. Limitaciones éticas y estructurales

### Estrategias de Mitigación Propuestas

La literatura analizada propone múltiples enfoques para reducir sesgos, aunque su implementación es incipiente (ver figura 4).<sup>(41,42)</sup> Una línea prometedora es el *fairness-aware machine learning*, que incorpora métricas de equidad durante el entrenamiento de modelos.<sup>(43)</sup> Técnicas como *reweighting* y *adversarial debiasing* han mejorado la precisión en grupos subrepresentados.<sup>(44)</sup>

En el plano institucional, se destacan las auditorías algorítmicas obligatorias, que exigen evaluaciones de impacto en derechos fundamentales.<sup>(45,46)</sup> Asimismo, los marcos éticos existentes pueden adaptarse para enfrentar los nuevos desafíos que plantea la IA en el diagnóstico de salud mental, y se proporcionan recomendaciones prácticas para los profesionales de la salud.<sup>(47)</sup>

Sin embargo, en opinión del autor, persisten desafíos en la escalabilidad de estas soluciones y en la voluntad política para adoptarlas. La justicia algorítmica en salud mental exigirá, además de innovación técnica, una redistribución de poder en la producción de conocimiento médico.

Enfoques técnicos para reducir sesgos	Medidas institucionales y regulatorias	Desafíos pendientes y consideraciones críticas
<ul style="list-style-type: none"> <li>• <i>Fairness-aware machine learning</i>: integración de métricas de equidad en el entrenamiento de modelos.</li> <li>• Técnicas como <i>reweighting</i> y <i>adversarial debiasing</i> para mejorar precisión en grupos subrepresentados.</li> </ul>	<ul style="list-style-type: none"> <li>• Implementación de auditorías algorítmicas obligatorias con enfoque en derechos fundamentales.</li> <li>• Adaptación de marcos éticos existentes para abordar desafíos de la IA en salud mental.</li> <li>• Recomendaciones prácticas para profesionales del sector.</li> </ul>	<ul style="list-style-type: none"> <li>• Limitaciones en la escalabilidad de las soluciones propuestas.</li> <li>• Falta de voluntad política para adoptar medidas de mitigación.</li> <li>• Necesidad de redistribución de poder en la producción de conocimiento médico.</li> </ul>

Figura 4. Estrategias para la mitigación de sesgos

## DISCUSIÓN

Los hallazgos de esta revisión revelan que los sesgos en los sistemas de diagnóstico basados en IA no son meras fallas técnicas, sino manifestaciones de inequidades estructurales profundamente arraigadas en los datos y los diseños algorítmicos.<sup>(23,26)</sup> La evidencia muestra que estos sesgos operan de manera diferencial, lo que afecta con mayor severidad a grupos históricamente marginados y reproduce patrones de exclusión en el acceso a la salud mental.<sup>(48)</sup> Esto plantea un desafío ético urgente, la necesidad de reconocer que la objetividad de los algoritmos es, en realidad, un espejo de los prejuicios existentes en las sociedades que los crean.<sup>(25)</sup> La solución no puede limitarse a ajustes técnicos superficiales, sino que requiere un cuestionamiento fundamental sobre quiénes participan en el desarrollo de estas tecnologías y qué voces son sistemáticamente silenciadas en el proceso.

Un aspecto crítico que emerge de la literatura es la tensión entre el potencial democratizador de la IA y su capacidad para amplificar desigualdades.<sup>(49)</sup> Mientras que estas herramientas prometen expandir el acceso a diagnósticos en regiones con escasez de especialistas, su implementación sin salvaguardas adecuadas puede perpetuar formas de colonialismo digital, donde poblaciones vulnerables quedan sujetas a sistemas que no comprenden ni controlan.<sup>(31)</sup> Este problema se agrava por la comercialización de algoritmos opacos, cuyos mecanismos internos son inaccesibles tanto para los profesionales de la salud como para los pacientes.<sup>(12,33)</sup> La falta de transparencia no solo limita la rendición de cuentas, sino que también erosiona la confianza en intervenciones que, paradójicamente, buscan mejorar la atención en salud mental.<sup>(50)</sup>

Las estrategias de mitigación analizadas, aunque prometedoras, enfrentan barreras significativas en su aplicación real. Iniciativas como los marcos de *fairness-by-design* y las auditorías participativas representan avances importantes, pero chocan con intereses comerciales, limitaciones presupuestarias y la ausencia de marcos regulatorios robustos.<sup>(38,46)</sup> Además, persiste una desconexión entre las soluciones técnicas propuestas y las necesidades concretas de los contextos locales, particularmente en países de ingresos bajos y medios. Esto subraya la importancia de desarrollar enfoques glociales que combinen estándares internacionales con adaptaciones culturalmente situadas, lo que evita soluciones universales que ignoran las particularidades de los sistemas de salud y las comunidades a las que sirven.<sup>(37)</sup>

Finalmente, esta revisión destaca que la lucha contra los sesgos algorítmicos en salud mental exige una acción colectiva y multidisciplinaria. Investigadores, clínicos, legisladores y comunidades afectadas deben colaborar para crear ecosistemas de IA ética que prioricen la justicia sobre la eficiencia.<sup>(46,47)</sup> Esto implica no solo mejorar los algoritmos, sino también transformar las estructuras de poder que determinan qué conocimientos se validan y qué poblaciones se consideran prioritarias.<sup>(4)</sup> El camino hacia diagnósticos algorítmicos verdaderamente equitativos requerirá, sobre todo, reconocer que la tecnología en salud mental no es neutral: es un campo de batalla donde se disputan visiones contrapuestas sobre quién merece ser escuchado y qué formas de sufrimiento son legítimas.

## CONCLUSIONES

Esta revisión documental evidencia que los sesgos algorítmicos en los sistemas de diagnóstico de salud mental reproducen y amplifican las desigualdades estructurales, al afectar desproporcionadamente a poblaciones vulnerables. Si bien se han identificado estrategias técnicas y éticas para mitigar estos sesgos, su implementación efectiva requiere un enfoque multidisciplinar que combine avances en inteligencia artificial con políticas públicas inclusivas, participación comunitaria y marcos regulatorios robustos. Garantizar la

equidad en el diagnóstico automatizado no es solo un desafío tecnológico, sino un imperativo ético que exige transformar los sistemas de salud mental para que prioricen la justicia social sobre la eficiencia algorítmica.

La superación de estos desafíos demanda una colaboración activa entre desarrolladores, profesionales de la salud, legisladores y comunidades afectadas, con el fin de crear herramientas que no solo sean precisas, sino también culturalmente sensibles y socialmente responsables. Solo mediante este compromiso colectivo podrá lograrse el potencial transformador de la IA en salud mental, lo que asegura que su aplicación beneficie a todos los grupos por igual, sin exacerbar las inequidades existentes.

## REFERENCIAS BIBLIOGRÁFICAS

1. Graham S, Depp C, Lee E, Nebeker C, Tu X, Kim H, Jeste D. Artificial Intelligence for Mental Health and Mental Illnesses: an Overview. *Current Psychiatry Reports* 2019;21:116. <https://doi.org/10.1007/s11920-019-1094-0>
2. Grzenda A. Artificial Intelligence in Mental Health. In: *Convergence Mental Health*. Oxford University Press; 2021. <https://doi.org/10.1093/MED/9780197506271.003.0011>
3. Pérez Gamboa AJ, Gómez Cano CA, Sánchez Castillo V. Decision making in university contexts based on knowledge management systems. *Data and Metadata* 2022;1:92. <https://doi.org/10.56294/dm202292>
4. Ntoutsis E, Fafalios P, Gadiraju U, Iosifidis V, Nejdil W, Vidal M, et al. Bias in data-driven artificial intelligence systems—An introductory survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 2020;10(3):e1356. <https://doi.org/10.1002/widm.1356>
5. Peters U. Algorithmic Political Bias in Artificial Intelligence Systems. *Philosophy & Technology* 2022;35:25. <https://doi.org/10.1007/s13347-022-00512-8>
6. Cano CAG, Castillo VS. Mobbing en una institución de educación superior en Colombia. *Aglala* 2021;12(2):100-116. <https://dialnet.unirioja.es/servlet/articulo?codigo=8453105>
7. Straw I, Callison-Burch C. Artificial Intelligence in mental health and the biases of language based models. *PLoS ONE* 2020;15(12):e0240376. <https://doi.org/10.1371/journal.pone.0240376>
8. Timmons A, Duong J, Fiallo N, Lee T, Vo H, Ahle M, et al. A Call to Action on Assessing and Mitigating Bias in Artificial Intelligence Applications for Mental Health. *Perspectives on Psychological Science* 2022;18(5):1062-1096. <https://doi.org/10.1177/17456916221134490>
9. Park J, Arunachalam R, Silenzio V, Singh V. Fairness in Mobile Phone-Based Mental Health Assessment Algorithms: Exploratory Study. *JMIR Formative Research* 2022;6:e34366. <https://doi.org/10.2196/34366>
10. Estrada-Cely GE, Sánchez-Castillo V, Gómez-Cano CA. Bioética y desarrollo sostenible: entre el biocentrismo y el antropocentrismo y su sesgo economicista. *Cilo América* 2018;12(24):255-267. <http://dx.doi.org/10.21676/23897848.2818>
11. Embí PJ. Algorithmic vigilance—Advancing Methods to Analyze and Monitor Artificial Intelligence-Driven Health Care for Effectiveness and Equity. *JAMA Network Open* 2021;4(4):e214622. <https://doi.org/10.1001/jamanetworkopen.2021.4622>
12. Gooding P, Kariotis T. Ethics and Law in Research on Algorithmic and Data-Driven Technology in Mental Health Care: Scoping Review. *JMIR Mental Health* 2021;8(6):e24668. <https://doi.org/10.2196/24668>
13. Gómez-Cano CA, Sánchez V, Tovar G. Factores endógenos causantes de la permanencia irregular: una lectura desde el actuar docente. *Educación y Humanismo* 2018;20(35):96-112. <https://revistas.unisimon.edu.co/index.php/educacion/article/view/3030>
14. Schouler-Ocak M. Transcultural Aspect of Mental Health Care. *European Psychiatry* 2022;65(S1):S3. <https://doi.org/10.1192/j.eurpsy.2022.37>
15. Schick A, Rauschenberg C, Ader L, Daemen M, Wieland L, Paetzold I, et al. Novel digital methods for gathering intensive time series data in mental health research: scoping review of a rapidly evolving field.

Psychological Medicine 2023;53(1):55-65. <https://doi.org/10.1017/S0033291722003336>

16. Guayara Cuéllar CT, Millán Rojas EE, Gómez Cano CA. Diseño de un curso virtual de alfabetización digital para docentes de la Universidad de la Amazonia. *Revista científica* 2019;(34):34-48. <https://doi.org/10.14483/23448350.13314>

17. Gómez Cano CA, Sánchez Castillo V, Millán Rojas EE. Capitalismo y ética: una relación de tensiones. *Económicas CUC* 2019;40(2):31-42. <https://doi.org/10.17981/econcuc.40.2.2019.02>

18. Siddaway AP, Wood AM, Hedges LV. How to Do a Systematic Review: A Best Practice Guide for Conducting and Reporting Narrative Reviews, Meta-Analyses, and Meta-Syntheses. *Annual Review of Psychology* 2019;70:747-770. <https://doi.org/10.1146/annurev-psych-010418-102803>

19. Hiebl MRW. Sample Selection in Systematic Literature Reviews of Management Research. *Organizational Research Methods* 2023;26(2):229-261. <https://doi.org/10.1177/1094428120986851>

20. Pérez Gamboa AJ, García Acevedo Y, García Batán J. Proyecto de vida y proceso formativo universitario: un estudio exploratorio en la Universidad de Camagüey. *Transformación* 2019;15(3):280-296. [http://scielo.sld.cu/scielo.php?script=sci\\_arttext&pid=S2077-29552019000300280](http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S2077-29552019000300280)

21. Richter T, Fishbain B, Fruchter E, Richter-Levin G, Okon-Singer H. Machine learning-based diagnosis support system for differentiating between clinical anxiety and depression disorders. *Journal of Psychiatric Research* 2021;141:199-205. <https://doi.org/10.1016/j.jpsychires.2021.06.044>

22. Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. *Science* 2019;366(6464):447-453. <https://doi.org/10.1126/science.aax2342>

23. Rai T. Racial bias in health algorithms. *Science* 2019;366(6464):440. <https://doi.org/10.1126/science.366.6464.440-e>

24. Hooker K, Phibbs S, Irvin V, Mendez-Luck C, Doan L, Li T, et al. Depression Among Older Adults in the United States by Disaggregated Race and Ethnicity. *The Gerontologist* 2019;59(5):886-891. <https://doi.org/10.1093/geront/gny159>

25. Sedgewick F, Kerr-Gaffney J, Leppanen J, Tchanturia K. Anorexia nervosa, autism, and the ADOS: How appropriate is the new algorithm for identifying cases? *Frontiers in Psychiatry* 2019;10:507. <https://doi.org/10.3389/fpsy.2019.00507>

26. Berthonnet I. Where Exactly Does the Sexist Bias in the Official Measurement of Monetary Poverty in Europe Come From? *Review of Radical Political Economics* 2021;55(1):132-146. <https://doi.org/10.1177/0486613420981785>

27. Ciecierski-Holmes T, Singh R, Axt M, Brenner S, Barteit S. Artificial intelligence for strengthening healthcare systems in low- and middle-income countries: a systematic scoping review. *NPJ Digital Medicine* 2022;5:162. <https://doi.org/10.1038/s41746-022-00700-y>

28. Gomez Cano CA, Sánchez Castillo V, Clavijo Gallego TA. English teaching in undergraduate programs: A reading of the challenges at Uniamazonia from teaching practice. *Horizontes Pedagógicos* 2018;20(1):55-62. <https://doi.org/10.33881/0123-8264.hop.20107>

29. Buda T, Guerreiro J, Iglesias J, Castillo C, Smith O, Matic A. Foundations for fairness in digital health apps. *Frontiers in Digital Health* 2022;4:943514. <https://doi.org/10.3389/fdgth.2022.943514>

30. Xu J, Xiao Y, Wang W, Ning Y, Shenkman E, Bian J, Wang F. Algorithmic fairness in computational medicine. *EBioMedicine* 2022;84:104250. <https://doi.org/10.1016/j.ebiom.2022.104250>

31. Gloria K, Rastogi N, DeGroff S. Bias Impact Analysis of AI in Consumer Mobile Health Technologies: Legal, Technical, and Policy. *arXiv* 2022;abs/2209.05440. <https://doi.org/10.48550/arXiv.2209.05440>

32. Simpson E, Semaan B. For You, or For" You"? . Proceedings of the ACM on Human-Computer Interaction 2021;4(CSCW3):1-34. <https://doi.org/10.1145/3432951>
33. Oliva T, Antonialli D, Gomes A. Fighting Hate Speech, Silencing Drag Queens? Artificial Intelligence in Content Moderation and Risks to LGBTQ Voices Online. *Sexuality & Culture* 2021;25(2):700-732. <https://doi.org/10.1007/s12119-020-09790-w>
34. Wylie L, Van Meyel R, Harder H, Sukhera J, Luc C, Ganjavi H, Elfakhani M, Wardrop N. Assessing trauma in a transcultural context: challenges in mental health care with immigrants and refugees. *Public Health Reviews* 2018;39:22. <https://doi.org/10.1186/s40985-018-0102-y>
35. Huq S, Maskeliūnas R, Damaševičius R. Dialogue agents for artificial intelligence-based conversational systems for cognitively disabled: a systematic review. *Disability and Rehabilitation: Assistive Technology* 2022;19(5):1059-1078. <https://doi.org/10.1080/17483107.2022.2146768>
36. Lee-Cheong S, Amanullah S, Jardine M. New assistive technologies in dementia and mild cognitive impairment care: A PubMed review. *Asian Journal of Psychiatry* 2022;73:103135. <https://doi.org/10.1016/j.ajp.2022.103135>
37. Wangmo T, Lipps M, Kressig R, Ienca M. Ethical concerns with the use of intelligent assistive technology: findings from a qualitative study with professional stakeholders. *BMC Medical Ethics* 2019;20:98. <https://doi.org/10.1186/s12910-019-0437-z>
38. Pérez A, Echerri D, García Y. Life project as a category of Higher Education pedagogy: Approaching a grounded theory. *Transformación* 2021;17(3):542-563. <http://scielo.sld.cu/pdf/trf/v17n3/2077-2955-trf-17-03-542.pdf>
39. Bird K, Castleman B, Mabel Z, Song Y. Bringing Transparency to Predictive Analytics: A Systematic Comparison of Predictive Modeling Methods in Higher Education. *AERA Open* 2021;7:23328584211037630. <https://doi.org/10.1177/23328584211037630>
40. Gómez Cano CA, García Acevedo Y, Pérez Gamboa AJ. Intersection between health and entrepreneurship in the context of sustainable development. *Health Leadership and Quality of Life* 2022;1:89. <https://doi.org/10.56294/hl202289>
41. Norori N, Hu Q, Aellen F, Faraci F, Tzovara A. Addressing bias in big data and AI for health care: A call for open science. *Patterns* 2021;2(10):100347. <https://doi.org/10.1016/j.patter.2021.100347>
42. Fazelpour S, Danks D. Algorithmic bias: Senses, sources, solutions. *Philosophy Compass* 2021;16(8):e12760. <https://doi.org/10.1111/PHC3.12760>
43. Bird S, Kenthapadi K, Kıcıman E, Mitchell M. Fairness-Aware Machine Learning: Practical Challenges and Lessons Learned. Proceedings of the 12th ACM International Conference on Web Search and Data Mining 2019:834-835. <https://doi.org/10.1145/3289600.3291383>
44. Petrović A, Nikolić M, Radovanović S, Delibašić B, Jovanović M. FAIR: Fair Adversarial Instance Re-weighting. *Neurocomputing* 2022;476:14-37. <https://doi.org/10.1016/j.neucom.2021.12.082>
45. Mantelero A. AI and Big Data: A Blueprint for a Human Rights, Social and Ethical Impact Assessment. *Computer Law & Security Review* 2018;34(4):754-772. <https://doi.org/10.1016/J.CLSR.2018.05.017>
46. Yam J, Skorburg J. From human resources to human rights: Impact assessments for hiring algorithms. *Ethics and Information Technology* 2021;23(4):611-623. <https://doi.org/10.1007/s10676-021-09599-7>
47. Straw I. Ethical implications of emotion mining in medicine. *Health Policy and Technology* 2021;10(1):167-171. <https://doi.org/10.1016/j.hlpt.2020.11.006>
48. Walsh C, Chaudhry B, Dua P, Goodman K, Kaplan B, Kavuluru R, Solomonides A, Subbian V. Stigma, biomarkers, and algorithmic bias: recommendations for precision behavioral health with artificial intelligence.

JAMIA Open 2020;3(1):9-15. <https://doi.org/10.1093/jamiaopen/ooz054>

49. Koster R, Balaguer J, Tacchetti A, Weinstein A, Zhu T, Hauser O, et al. Human-centred mechanism design with Democratic AI. *Nature Human Behaviour* 2022;6(10):1398-1407. <https://doi.org/10.1038/s41562-022-01383-x>

50. D'Alfonso S. AI in mental health. *Current Opinion in Psychology* 2020;36:112-117. <https://doi.org/10.1016/j.copsyc.2020.04.005>

### **FINANCIACIÓN**

Los autores no recibieron financiación para el desarrollo de la presente investigación.

### **CONFLICTO DE INTERESES**

Los autores declaran que no existe conflicto de intereses.

### **CONTRIBUCIÓN DE AUTORÍA**

*Conceptualización:* Ariadna Matos Matos.

*Curación de datos:* Ariadna Matos Matos.

*Análisis formal:* Ariadna Matos Matos.

*Investigación:* Ariadna Matos Matos.

*Metodología:* Ariadna Matos Matos.

*Administración del proyecto:* Ariadna Matos Matos.

*Recursos:* Ariadna Matos Matos.

*Software:* Ariadna Matos Matos.

*Supervisión:* Ariadna Matos Matos.

*Validación:* Ariadna Matos Matos.

*Visualización:* Ariadna Matos Matos.

*Redacción - borrador original:* Ariadna Matos Matos.

*Redacción - revisión y edición:* Ariadna Matos Matos.