



POSICIONAMIENTO

Inteligencia Artificial y Sesgo de Género: Desafíos y Oportunidades para la Equidad en Salud

Posicionamiento del Grupo de Trabajo de Género, Diversidad y Salud de la Sociedad Española de Epidemiología

Autoría: Blanca Obón Azuara, Christian Gil Borrelli, Isabel Gutiérrez Cía, Marta Puig-García, Carmen Vives-Cases

Introducción

La integración de los sistemas de inteligencia artificial (IA) en la práctica clínica presenta una oportunidad sin precedentes para mejorar la salud, pero también plantea importantes desafíos para la equidad de género, con consecuencias directas y significativas para la salud de las mujeres [1]. La IA influye en decisiones médicas en cientos de hospitales a través de herramientas de diagnóstico y pronóstico [2,3]. Sin embargo, es fundamental asegurar que estos algoritmos se desarrollen y apliquen de manera equitativa. Estudios recientes muestran que, cuando no se gestionan adecuadamente, surgen sesgos de género. Se ha observado que algoritmos para enfermedad hepática pueden tener una menor tasa de acierto en mujeres en comparación con los hombres [4], y que modelos de lenguaje como GPT-4 podrían evaluar de manera diferente el riesgo cardiovascular en mujeres si presentan comorbilidades psiquiátricas [5]. Estos hallazgos subrayan la necesidad de un enfoque de género proactivo para garantizar que la IA se convierta en una herramienta para reducir, y no para perpetuar, las disparidades existentes.

Este desafío se origina por la falta de una perspectiva de género en el “ciclo de vida” de la IA, desde la recopilación de datos hasta su implementación. Como resultado, se han identificado riesgos de diagnósticos médicos menos precisos en las mujeres, así como sesgos en otros ámbitos como el laboral, el judicial o el financiero [6,7]. Desde una perspectiva epidemiológica y de salud pública, el sesgo de género en la IA representa un reto fundamental para la validez de la vigilancia epidemiológica, la efectividad de las intervenciones preventivas y la equidad en la distribución de recursos sanitarios [8,9,10]. Este documento de posicionamiento consolida la evidencia científica más reciente sobre este tema y propone un plan de acción constructivo con el objetivo de promover una IA justa, segura y que contribuya activamente a la equidad de género en la salud y en la sociedad.

1. El Origen del Desafío: Equidad Algorítmica y Representatividad de los Datos

Uno de los principales desafíos para la equidad de género en la IA reside en los datos históricos con los que se entrenan los algoritmos, que reproducen los sesgos sociales preexistentes. En el ámbito de la salud, ha sido una práctica extendida el uso del “varón blanco estándar” como perfil clínico de referencia. Esta aproximación metodológica puede invisibilizar sistemáticamente los síntomas, las necesidades y las particularidades fisiológicas de las mujeres y de las personas con identidades de género diversas [11,12]. Esta situación



representa un importante desafío estructural que puede comprometer la validez clínica y epidemiológica de los modelos de IA. Cuando un algoritmo no se entrena con datos que representan adecuadamente a toda la población, su capacidad predictiva es menos precisa para los grupos infrarrepresentados, lo que puede conducir a estimaciones menos exactas de prevalencia, incidencia y riesgo [13,14].

Investigaciones sobre la aplicación de IA y aprendizaje automático (*machine learning*) en el sector salud revelan que hasta un 31% de los estudios omiten el género en sus conjuntos de datos [15]. Esta falta de diversidad en los datos de entrenamiento puede reproducir patrones de exclusión que se trasladan a la toma de decisiones algorítmicas. Desde la perspectiva de la salud pública, esta “invisibilidad de datos” no solo representa un reto para la atención clínica individual, sino que también puede distorsionar la planificación de políticas sanitarias, la asignación de recursos y el desarrollo de estrategias de prevención [16,17].

Este fenómeno se extiende más allá del ámbito sanitario. Los grandes modelos de lenguaje (LLM) tienden a asociar de manera sistemática una mayor responsabilidad en la crianza y los cuidados a las mujeres [18]. Este sesgo no se debe a una programación explícita, sino a las asociaciones contextuales aprendidas de los corpus de texto con los que son entrenados, que reflejan normas sociales tradicionales. Esto subraya la importancia de considerar las implicaciones de la IA en la salud mental de las mujeres, la carga de cuidados no remunerados y la equidad en la conciliación de la vida laboral y familiar.

2. Impacto Multisectorial del Sesgo de Género en la IA

El sesgo de género algorítmico es un desafío transversal que se manifiesta en múltiples dominios, pudiendo crear un ciclo de retroalimentación que perpetúe las desigualdades sociales y afecte la salud de las mujeres. A continuación, se detallan los impactos más relevantes en diferentes sectores:

2.1 Salud: Hacia un Diagnóstico Preciso y una Atención Adecuada

La comunidad científica ha señalado que algunos sistemas de IA diseñados para el apoyo al diagnóstico clínico pueden mostrar un rendimiento inferior en mujeres, especialmente en patologías como las enfermedades cardiovasculares. Esto puede deberse a que los modelos son entrenados predominantemente con datos de pacientes masculinos, que puede llevar a la omisión de presentaciones sintomáticas en mujeres, con el consiguiente riesgo de retrasos en la atención y un aumento de la morbilidad evitable [11,19,20].

Como consecuencia, si las recomendaciones generadas por la IA son sesgadas, existe el riesgo de dejar a las mujeres sin la atención necesaria y adecuada [21,22]. Si un algoritmo subestima el riesgo de una mujer o no reconoce sus síntomas, la recomendación resultante podría ser la de no actuar o actuar con menor urgencia. Un estudio en el *Journal of Medical Internet Research* observó que GPT-4 cambiaba su evaluación del riesgo de enfermedad coronaria para una mujer si se le añadía un diagnóstico de salud mental, mientras que no lo hacía para un hombre con el mismo perfil clínico [5]. Este tipo de sesgo puede traducirse en una menor probabilidad de que los/las médico/as soliciten pruebas diagnósticas o inicien un tratamiento preventivo.



Desde una perspectiva de salud pública y epidemiología, esta situación plantea importantes retos:

- Compromete la validez de los sistemas de vigilancia epidemiológica basados en IA, pudiendo generar subestimaciones de la carga de enfermedad en mujeres.
- Afecta la efectividad de las intervenciones preventivas, ya que los algoritmos de estratificación de riesgo podrían no identificar correctamente a las mujeres en situación de riesgo elevado.
- Puede generar inequidades en la distribución de recursos sanitarios, al asignar prioridades diagnósticas y terapéuticas basadas en modelos sesgados [23,24].

Además, los sistemas sesgados corren el riesgo de influir en la elaboración de guías de práctica clínica. Si las organizaciones sanitarias adoptan herramientas de IA para estandarizar protocolos, las guías resultantes podrían integrar y automatizar estas desigualdades a gran escala, consolidando una práctica médica de menor calidad para las mujeres de manera estructural [25,26].

2.2 Empleo, Justicia y Finanzas: Determinantes Sociales de la Salud

El impacto del sesgo algorítmico se extiende a otros ámbitos que actúan como determinantes sociales de la salud:

- Empleo: Se ha observado que algunos algoritmos de selección de personal tienden a penalizar currículums que contienen términos asociados a perfiles femeninos (p. ej., “colaborativa”, “cuidadora”). Asimismo, los modelos que predicen el desempeño laboral tienden a valorar atributos tradicionalmente masculinizados, lo que podría reforzar la brecha de género en la promoción profesional [27,28]. Esta discriminación algorítmica puede contribuir a la precariedad laboral, la brecha salarial y el estrés psicosocial, factores que impactan negativamente en la salud de las mujeres [29].
- Justicia: Los sistemas de reconocimiento facial han mostrado tasas de error significativamente más altas en las mujeres racializadas, lo que ha generado preocupación por detenciones injustificadas. A su vez, los algoritmos que evalúan el riesgo de reincidencia penal a menudo utilizan variables indirectas (código postal, nivel de ingresos) que pueden penalizar a colectivos históricamente excluidos, afectando su salud mental y su acceso a servicios [30,31,32].
- Finanzas: Existe el riesgo de que los modelos de *scoring* crediticio repliquen discriminaciones históricas del sistema financiero, asignando peores condiciones o denegando préstamos a mujeres, especialmente a emprendedoras o a aquellas fuera del mercado laboral formal. Esta exclusión financiera puede limitar sus oportunidades económicas, contribuyendo a ciclos de pobreza y precariedad [33].

2.3 Representación Profesional: Estereotipos en la Medicina

La IA generativa, utilizada en la creación de contenidos para formación o divulgación médica, puede reforzar estereotipos de género. Un estudio reciente demostró que estos sistemas representan a los hombres como médicos especialistas en un 82% de las imágenes generadas,



a pesar de que solo ocupan el 47% de dichos cargos en la realidad [34]. Las mujeres, por su parte, son sistemáticamente excluidas de especialidades como urología o traumatología y sobrerrepresentadas en dermatología u obstetricia [34].

Esta representación visual sesgada responde a lo que se ha identificado como un “techo de cristal algorítmico”, donde los modelos de lenguaje y visión tienden a simplificar la complejidad social, seleccionando los rasgos más prototípicos y prejuiciosos para representar el prestigio profesional [35]

Al hacerlo, la IA no solo replica la desigualdad, sino que la intensifica, creando un entorno digital donde la autoridad médica se vuelve por defecto visualmente masculina.

Algunas implicaciones de este fenómeno incluyen:

- Construcción simbólica del poder: La IA tiende a asociar la competencia técnica y la alta especialización con la masculinidad, mientras vincula la femineidad a roles de soporte o estéticos, reforzando construcciones culturales que perpetúan la jerarquía de género en las organizaciones sanitarias [36].
- Impacto en la identidad profesional: La representación sesgada puede limitar las aspiraciones profesionales de las nuevas generaciones, enviando un mensaje implícito sobre qué perfiles “encajan” en determinadas especialidades, lo que puede provocar consecuencias sobre la composición y diversidad de la fuerza laboral sanitaria.
- Validación de sesgos en salud pública: Si las instituciones utilizan estas imágenes en materiales educativos o campañas de promoción de la salud, pueden seguir fortaleciendo estas representaciones estereotipadas y poco diversas, validándolas institucionalmente.

3. Repercusiones para la Epidemiología y la Salud Pública

El sesgo de género en la IA plantea desafíos específicos y relevantes para la práctica de la epidemiología y la salud pública, que van más allá de los efectos clínicos individuales:

- 1 Compromiso de la Validez de la Vigilancia Epidemiológica: Los sistemas de vigilancia basados en IA pueden generar estimaciones sesgadas de prevalencia, incidencia y mortalidad en mujeres si los algoritmos de detección y clasificación de casos están entrenados con datos no representativos. Esto puede comprometer la capacidad de los sistemas de salud pública para identificar tendencias, detectar brotes y evaluar la carga de enfermedad en las mujeres de manera precisa [37,38].
- 2 Inequidad en la Asignación de Recursos Sanitarios: Los algoritmos de priorización y asignación de recursos pueden perpetuar o amplificar inequidades de género. Por ejemplo, algoritmos que predicen el riesgo de hospitalización o la necesidad de cuidados intensivos podrían subestimar las necesidades de las mujeres, resultando en un menor acceso a servicios de medicina intensiva y en una distribución injusta de los recursos disponibles [26,39].
- 3 Inefectividad de las Intervenciones Preventivas: Los modelos de estratificación de riesgo utilizados para dirigir intervenciones preventivas (como el cribado de cáncer o la prevención cardiovascular) pueden fallar en la identificación correcta de mujeres en



riesgo elevado. Esto compromete la efectividad y la equidad de las estrategias de prevención, dejando a una parte de la población desprotegida [40,41].

- 4 Perpetuación de los Determinantes Sociales de la Salud: El sesgo de género en la IA aplicada a sectores como el empleo, las finanzas y la justicia puede contribuir a perpetuar determinantes sociales de la salud como la precariedad laboral, la pobreza o la discriminación. Todos ellos son factores de riesgo establecidos para múltiples enfermedades crónicas y problemas de salud mental en mujeres [29,42].
- 5 Erosión de la Confianza en los Sistemas de Salud: La experiencia repetida de diagnósticos menos precisos, retrasos en la atención y la falta de reconocimiento de sus síntomas puede erosionar la confianza de las mujeres en los sistemas de salud. Esto tiene consecuencias directas en la adherencia terapéutica, la utilización de servicios preventivos y la participación en programas de salud pública [43].

4. Síntesis de Consecuencias del Sesgo de Género en la IA

La siguiente tabla sintetiza las principales consecuencias del sesgo de género algorítmico en diferentes sectores y su impacto directo en la equidad y la salud pública:

Sector	Consecuencia Directa	Impacto en la Equidad de Género y la Salud Pública
Salud	Diagnósticos menos precisos, falsos negativos y guías clínicas influenciadas por sesgos.	Riesgo de retraso en el diagnóstico y tratamiento, con el consecuente aumento de la morbilidad evitable en mujeres; compromiso de la validez de la vigilancia epidemiológica e ineficacia de las intervenciones preventivas [19,20,25,26].
	Desigualdad en el acceso a tratamientos y recursos sanitarios.	Subestimación de la gravedad de enfermedades en mujeres, restricción del acceso a tratamientos y perpetuación de inequidades en la asignación de recursos [21,22,37].
Empleo	Discriminación en la contratación y promoción profesional.	Los algoritmos pueden favorecer perfiles masculinos, contribuyendo a la segregación ocupacional, limitando oportunidades, reforzando la brecha salarial e impactando en los determinantes sociales de la salud [27,28,29].



Sector	Consecuencia Directa	Impacto en la Equidad de Género y la Salud Pública
Justicia	Identificación injusta y sesgos en la evaluación del riesgo penal.	Altas tasas de error en reconocimiento facial para mujeres racializadas, que pueden conducir a detenciones injustas; discriminación en el sistema penal con impacto en la salud mental y el acceso a servicios [30,31,32].
Finanzas	Denegación de crédito o imposición de condiciones desfavorables.	Obstáculos al acceso a capital y al emprendimiento femenino, lo que puede perpetuar la pobreza y la precariedad como determinantes sociales de la salud [33].
Social	Refuerzo de estereotipos de género y roles de cuidado.	Normalización de roles tradicionales, perpetuación de la desigualdad, aumento de la carga de cuidados no remunerados e impacto en la salud mental y la conciliación laboral-familiar [18,34].

5. Plan de Acción: Hacia una Inteligencia Artificial con Perspectiva de Género

Para abordar estos desafíos y avanzar hacia un ecosistema de IA que promueva la equidad en salud, el Grupo de Trabajo de Género, Diversidad y Salud de la SEE propone un plan de acción integral y constructivo, centrado en cinco ejes estratégicos:

Paso 1: Auditoría y Reestructuración de Conjuntos de Datos Clínicos

Objetivo:

Corregir la “invisibilidad de datos” asegurando una representación equitativa de género, raza, edad y otras variables sociodemográficas en los datos de entrenamiento.

Es imperativo revisar la composición de los datasets para corregir la subrepresentación histórica de mujeres, así como de las personas trans y de los colectivos racializados. Las estrategias deben incluir la recolección activa de datos clínicos específicos de mujeres, el uso de técnicas de sobremuestreo (oversampling) y ponderación estadística, y el empleo de variables interseccionales [13,17,20]. Se debe llevar a cabo una auditoría rigurosa de todos los conjuntos de datos utilizados para entrenar modelos de diagnóstico, triaje y vigilancia epidemiológica, identificando vacíos de representatividad y diseñando estrategias correctivas basadas en evidencia epidemiológica [44,45]. Desde una perspectiva de salud pública, es fundamental que los datos de entrenamiento reflejen la diversidad poblacional real para garantizar la validez externa de los modelos [46].



Paso 2: Desarrollo de Modelos con Sensibilidad de Género y Validación Interseccional

Objetivo:

Exigir que los modelos de IA demuestren un rendimiento comparable y equitativo en todos los subgrupos demográficos relevantes.

La validación de algoritmos debe ir más allá de la precisión global e incorporar métricas de equidad, como la sensibilidad, especificidad y los valores predictivos desglosados por subgrupo demográfico. La aprobación regulatoria de un modelo clínico debería estar condicionada a que su rendimiento sea comparable entre hombres, mujeres y otros grupos de interés [11,12,13]. Los desarrolladores deben validar el desempeño de sus modelos en subgrupos diferenciados por sexo, género, raza y otros ejes de desigualdad, garantizando que la IA no solo sea precisa en general, sino también justa en su aplicación clínica y epidemiológica [47,48]. Se debería exigir la incorporación de la perspectiva interseccional en todas las fases del desarrollo algorítmico para evaluar cómo el género interactúa con otras identidades (raza, clase, edad, discapacidad) y evitar discriminaciones múltiples [49].

Paso 3: Implementación de Protocolos de Supervisión Humana Crítica y Alertas de Sesgo

Objetivo:

Establecer un mecanismo de seguridad clínica donde la decisión algorítmica sea siempre revisada por un profesional capacitado.

Se debe integrar la supervisión humana en todo proceso algorítmico crítico, especialmente en decisiones que afecten directamente a la salud o a los derechos fundamentales. Todos los sistemas de IA deben formar parte de protocolos clínicos que exijan la revisión por parte de los/las profesionales ante situaciones de alto riesgo, casos clínicamente atípicos o cuando el paciente pertenezca a grupos subrepresentados en los datos [11]. Además, los sistemas deben incluir alertas automáticas que notifiquen sobre sesgos conocidos o posibles fallos en poblaciones infrarrepresentadas, forzando los/las profesionales sanitarios/as a una revisión manual y consciente del diagnóstico [5]. Desde una perspectiva de salud pública, es fundamental que los/las profesionales de la epidemiología y la salud pública tengan formación en la identificación y corrección de sesgos algorítmicos [50].

Paso 4: Establecimiento de Estándares Regulatorios de Equidad de Datos

Objetivo:

Crear un marco legal y ético que haga obligatoria la documentación y el cumplimiento de la diversidad de datos para la certificación de la IA médica.



Las agencias reguladoras, como la Agencia Europea de Medicamentos (EMA) o la FDA, deben establecer requisitos explícitos de equidad algorítmica. Esto incluye la obligación de presentar informes de impacto por género, raza y clase como parte del proceso de certificación [17,51]. La certificación debe estar condicionada a la demostración de que el modelo no amplifica las disparidades de salud existentes, priorizando la seguridad del paciente y la equidad sobre la mera eficiencia algorítmica. El principio rector debe ser que ninguna herramienta basada en IA pueda ser aprobada si amplifica disparidades estructurales en salud, tal y como señala la Organización Mundial de la Salud (OMS) [51]. Además, se deben implementar marcos regulatorios que obliguen a la auditoría externa e independiente de los conjuntos de datos y de los procesos de decisión algorítmica para detectar y corregir sesgos de manera proactiva [52,53].

Paso 5: Creación de un Sistema de Monitorización Post-Implementación y Retroalimentación Clínica

Objetivo:

Asegurar un ciclo de mejora continua que detecte y corrija el sesgo residual en tiempo real, basándose en resultados clínicos del mundo real.

Los modelos deben estar sujetos a una vigilancia activa tras su implementación. El monitoreo clínico en entornos reales permite identificar sesgos residuales, los cuales deben retroalimentar mejoras iterativas de los algoritmos [13,17,51]. Es imprescindible implementar sistemas de monitoreo continuo del rendimiento del modelo en poblaciones clínicas reales, acompañado de canales de retroalimentación que permitan a los profesionales reportar errores o comportamientos anómalos de la IA. Los datos de estos fallos deben ser utilizados para ajustar y reentrenar los modelos de forma iterativa [54]. Desde una perspectiva de salud pública, es fundamental que los sistemas de monitoreo post-implementación incluyan indicadores de equidad de género, como diferencias en tasas de falsos negativos, acceso a servicios y resultados de salud por sexo y género [55].

6. Recomendaciones Finales del Grupo de Trabajo

El desarrollo ético y justo de la IA exige un compromiso explícito con la equidad de género y la justicia social. Desde el Grupo de Trabajo de Género, Diversidad y Salud de la Sociedad Española de Epidemiología, instamos a:

- 1** Incorporar enfoques interseccionales desde la etapa de diseño algorítmico, reconociendo que las desigualdades de género en salud son el resultado de la interacción de múltiples ejes de opresión (sexo, género, raza, clase, edad, discapacidad).



- 2 Realizar auditorías independientes y públicas de datos y resultados, con participación de expertos en epidemiología, salud pública, ética y estudios de género, y con transparencia en la publicación de resultados.
- 3 Garantizar la diversidad interdisciplinaria en los equipos de desarrollo tecnológico, promoviendo activamente la participación de mujeres y minorías de género, así como de expertas en ética, ciencias sociales y salud de género.
- 4 Aplicar regulaciones que prioricen la equidad sobre la eficiencia algorítmica, estableciendo que ninguna herramienta de IA debe ser aprobada si amplifica desigualdades estructurales en salud, tal como señala la OMS.
- 5 Fortalecer la formación de los profesionales de la salud pública y la epidemiología en la identificación, evaluación y corrección de sesgos algorítmicos, y en el uso crítico de herramientas de IA en la práctica epidemiológica y la planificación sanitaria.
- 6 Promover la investigación epidemiológica sobre el impacto de la IA en las desigualdades de género en salud, incluyendo estudios de evaluación de impacto, análisis de equidad y monitoreo de indicadores de salud desagregados por sexo y género.
- 7 Fomentar la participación ciudadana y de organizaciones de mujeres en el diseño, evaluación y monitoreo de sistemas de IA aplicados a la salud, garantizando que las voces de las personas afectadas sean escuchadas y tenidas en cuenta en las decisiones tecnológicas.

La inteligencia artificial es una herramienta poderosa con un gran potencial para transformar la salud. Pero su desarrollo debe ser inclusivo, equitativo y basado en evidencia epidemiológica sólida para que no se convierta en una herramienta más para la perpetuación de la desigualdad. Desde la epidemiología y la salud pública, tenemos la responsabilidad de guiar el desarrollo de las tecnologías emergentes para que contribuyan a reducir —y no a amplificar— las brechas de género en salud.

Referencias

1. Kaushal A, Altman R, Langlotz C. Underdiagnosis bias of artificial intelligence algorithms in radiology. *Nat Med.* 2021;27(12):2098-2100.
2. Wong A, Otles E, Donnelly JP, et al. External validation of a widely implemented proprietary sepsis prediction model in hospitalized patients. *JAMA Intern Med.* 2021;181(8):1065-1070.
3. Chen J, Miao C. DeepSeek Deployed in 90 Chinese Tertiary Hospitals: How Artificial Intelligence Is Transforming Clinical Practice. *J Med Syst.* 2025;49(1):17.
4. Straw I, Wu Z. Investigating for bias in healthcare algorithms: a sex-stratified analysis of supervised machine learning models in liver disease prediction. *BMJ Health Care Inform.* 2022;29(1):e100457.
5. Mellino S, Giarratano Y, Nicolò A, et al. Gender Bias in AI's Perception of Cardiovascular Risk. *J Med Internet Res.* 2024;26:e54242.
6. Fletcher RR, Nakeshimana A, Olubeko O, et al. Addressing fairness, bias, and appropriate use of artificial intelligence and machine learning in global health. *Front Artif Intell.* 2021;4:561802.



7. Joshi A, Pradhan B, Mittal P. Big data and AI for gender equality in health: bias is a big challenge. *Front Big Data*. 2024;7:1436019.
8. Celi LA, Cellini J, Charpignon ML, et al. Sources of bias in artificial intelligence that perpetuate healthcare disparities—A global review. *PLoS Digit Health*. 2022;1(3):e0000022.
9. Hernández-Aguado I, Chilet-Rosell E. *Epidemiología y políticas de salud pública*. *Gac Sanit*. 2020;34(Supl 1):1-3.
10. Bhopal RS. *Concepts of Epidemiology: Integrating the ideas, theories, principles, and methods of epidemiology*. 3rd ed. Oxford: Oxford University Press; 2016.
11. Cirillo D, Caturana-Solarz S, Morey C, et al. Sex and gender differences and biases in artificial intelligence for biomedicine and healthcare. *NPJ Digit Med*. 2020;3:81.
12. Bearman M, Ajjawi R, Kenealy S, et al. Artificial intelligence and gender equity. *Lancet Digit Health*. 2025;7(1):e15–e23.
13. Joseph J, Nair SC, Iyer S, Nath LR. Algorithmic bias in public health AI: a silent threat to equity in low-resource settings. *Front Public Health*. 2025;13:1643180.
14. Silva PC, Luz Filho SS, Prudêncio RBC, Barros de Almeida Neto A. Evaluating gender bias in ML-based clinical risk prediction models. *J Biomed Inform*. 2024;155:104102.
15. Otokiti AU, Olatosi B, Osuagwu CC, et al. Gender and racial bias unveiled: clinical artificial intelligence. *Oxf Open Digit Health*. 2025;3(1):oqaf027.
16. UN Women. *Cómo la inteligencia artificial refuerza los sesgos de género y qué podemos hacer al respecto* [Internet]. 2025 [citado 14 feb 2026]. Disponible en: <https://www.unwomen.org/es/noticias/entrevista/2025/02/como-la-inteligencia-artificial-refuerza-los-sesgos-de-genero-y-que-podemos-hacer-al-respecto>
17. D'Ignazio C, Klein LF. *Data Feminism*. Cambridge, MA: MIT Press; 2020.
18. Xiu Y, Xu Y, Yang X, Wang Y, Yip PSF, Lau JTF. Large language models reproduce gender bias in parenting. *PLoS One*. 2025;20(7):e0335706.
19. Criado Perez C. *Invisible Women: Data Bias in a World Designed for Men*. New York: Abrams; 2019.
20. Al Hamid A, Beckett R, Wilson M, et al. Gender bias in diagnosis, prevention, and treatment of cardiovascular diseases: a systematic review. *Cureus*. 2024;16(5):e60354.
21. Mittermaier M, Raza MM, Kvedar JC. Bias in AI-based models for medical applications. *NPJ Digit Med*. 2023;6(1):128.
22. Cross JL, Becker J, Emerson E, Gorman DR. Bias in medical AI: Implications for clinical decision-making. *EBioMedicine*. 2024;105:105277.
23. Lee MS, Floridi L, Singh J. Towards gender equity in artificial intelligence and machine learning in dermatology. *J Am Acad Dermatol*. 2021;85(6):1553-1555.
24. Isaksson A, Wallgren P, Gustafsson J, Karlsson D. Mitigation measures for addressing gender bias in artificial intelligence within healthcare. *AI Soc*. 2025.
25. American College of Physicians. Artificial intelligence in the provision of health care: an American College of Physicians policy position paper. *Ann Intern Med*. 2024;177(6):826-833.
26. 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.
27. Ho JQH. Gender biases within Artificial Intelligence and ChatGPT. *Safety Sci*. 2025;183:106428.
28. Forbes. *Women Who Use AI At Work Face A Predictable Competence Penalty* [Internet]. 2025 [citado 11 feb 2026]. Disponible en: <https://www.forbes.com>



29. Marmot M, Allen J, Goldblatt P, Herd E, Morrison J. Build Back Fairer: The COVID-19 Marmot Review. London: Institute of Health Equity; 2020.
30. ACLU. Biased Technology: The Automated Discrimination of Facial Recognition [Internet]. [citado 11 feb 2026]. Disponible en: <https://www.aclu.org>
31. Gentzel M, Garibay I, Manoharan S. Biased Face Recognition Technology Used by Government. *PMC*. 2021;8683193.
32. Krištofík A, Štefková Z. Bias in AI (Supported) Decision Making: Old Problems, New Challenges. *IAC J Comp Adm*. 2025;1(1):45-62.
33. Women's World Banking. Algorithmic Bias, Financial Inclusion, and Gender [Internet]. 2021 [citado 11 feb 2026]. Disponible en: <https://www.womensworldbanking.org>
34. Hartley S, Fisher J. Gender bias in AI-generated images of medical specialists: a comparison with real-world recruitment data. *Clin Teach*. 2025;22(4):e12797.
35. Sun L, Wei M, Sun Y, Suh YJ, Shen L. Smiling women pitching down: auditing representational and presentational gender biases in image-generative AI. *J Comput Mediat Commun*. 2024;29(1):zmad045. doi: 10.1093/jcmc/zmad045
36. Gorska AM, Jemielniak D. The invisible women: uncovering gender bias in AI-generated images of professionals. *Feminist Media Stud*. 2023;23(8):4370–4375. doi: 10.1080/14680777.2023.2241526
37. Redshaw M, Wynter K. Machine learning in women's health: a narrative review. *J Perinat Med*. 2022;50(9):1181-1191.
38. Muley A, Cheemalapati S, Muzumdar P, Deshpande S. The Role of AI in Epidemiological Research: Applications, Benefits, and Risks in Modern Public Health. *Int J Med Health Sci*. 2025;14(1):23-35.
39. Chin MH, Afsar-Manesh N, Bierman AS, et al. Guiding principles to address the impact of algorithm bias on racial and ethnic disparities in health and health care. *JAMA Netw Open*. 2023;6(12):e2345050.
40. Liu M, Caputi TL, Dredze M, Kesselheim AS, Ayers JW. A scoping review and evidence gap analysis of clinical AI fairness. *NPJ Digit Med*. 2025;8(1):167.
41. Lee B, Shinohara RT, Kuo PH, Patel MS. Early Recalls and Clinical Validation Gaps in Artificial Intelligence Medical Devices. *JAMA Health Forum*. 2025;6(5):e250782.
42. Solar O, Irwin A. A conceptual framework for action on the social determinants of health. *Social Determinants of Health Discussion Paper 2 (Policy and Practice)*. Geneva: World Health Organization; 2010.
43. Mihan A, Pandey A, Van Spall HGC. Artificial intelligence bias in the prediction and detection of cardiovascular disease. *npj Cardiovasc Health*. 2024;1:31.
44. Vyas DA, Eisenstein LG, Jones DS. Hidden in Plain Sight — Reconsidering the Use of Race Correction in Clinical Algorithms. *N Engl J Med*. 2020;383(9):874-882.
45. Hankinson JL, Odencrantz JR, Fedan KB. Spirometric reference values from a sample of the general U.S. population. *Am J Respir Crit Care Med*. 1999;159(1):179-187.
46. Rajkomar A, Hardt M, Howell MD, Corrado G, Chin MH. Ensuring Fairness in Machine Learning to Advance Health Equity. *Ann Intern Med*. 2018;169(12):866-872.
47. Gianfrancesco MA, Tamang S, Yazdany J, Schmajuk G. Potential Biases in Machine Learning Algorithms Using Electronic Health Record Data. *JAMA Intern Med*. 2018;178(11):1544-1547.
48. Bowleg L. The problem with the phrase women and minorities: intersectionality—an important theoretical framework for public health. *Am J Public Health*. 2012;102(7):1267-1273.



49. Habib AR, Gross CP. FDA regulations of AI-driven clinical decision support devices fall short. *JAMA Intern Med.* 2023;183(9):896-897.
50. World Health Organization. Ethics and governance of artificial intelligence for health: WHO guidance. Geneva: WHO; 2021.
51. Raji ID, Smart A, White RN, et al. Closing the AI accountability gap: defining an end-to-end framework for internal algorithmic auditing. In: *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency.* New York: ACM; 2020. p. 33-44.
52. Mittelstadt B. Principles alone cannot guarantee ethical AI. *Nat Mach Intell.* 2019;1(11):501-507.
53. Sendak MP, Gao M, Brajer N, Balu S. Presenting machine learning model information to clinical end users with model facts labels. *NPJ Digit Med.* 2020;3:41.
54. Braveman P, Gruskin S. Defining equity in health. *J Epidemiol Community Health.* 2003;57(4):254-258.