

Uso de la inteligencia artificial para prevenir agresiones a profesionales de la salud

Antonio J. Moreno-Moreno,¹ Juan J. García-Iglesias² y Juan Gómez-Salgado^{2,3*}

¹Departamento de Tecnologías de la Información. Universidad de Huelva, Huelva, España; ²Departamento de Sociología, Trabajo Social y Salud Pública, Facultad de Ciencias del Trabajo, Universidad de Huelva, Huelva, España; ³Programa de Posgrado en Seguridad y Salud, Universidad Espíritu Santo. Guayaquil, Ecuador

Resumen

El aumento de agresiones contra profesionales de la salud es un problema de salud pública y laboral que afecta el bienestar del personal y la calidad de la atención. La violencia en este sector incluye agresiones físicas, verbales y psicológicas, representando un riesgo significativo. Se distinguen cuatro tipos principales de violencia laboral: externa sin relación previa (Tipo I), de pacientes hacia profesionales (Tipo II, la más frecuente), interna o institucional (Tipo III) y personal (Tipo IV). Este problema es global, con una tendencia creciente y un subregistro considerable. Sus consecuencias afectan a nivel individual (burnout, ansiedad, depresión), institucional (absentismo, rotación de personal) y en la calidad de la atención. La inteligencia artificial (IA) se presenta como una herramienta prometedora para prevenir y mitigar estas agresiones. Sus aplicaciones incluyen vigilancia y monitoreo, mejora en la comunicación, optimización del trabajo, formación del personal y análisis predictivo de pacientes agresivos. No obstante, su implementación plantea desafíos éticos como la privacidad, el sesgo, la fiabilidad de predicciones y el riesgo de deshumanización. Para garantizar su uso seguro y equitativo, es esencial abordar estas cuestiones bajo supervisión humana y complementar la tecnología con estrategias organizativas y educativas.

PALABRAS CLAVE: Inteligencia artificial. Agresiones. Profesionales de la salud. Prevención. Violencia laboral. Subregistro.

Use of artificial intelligence to prevent aggressions against health professionals

Abstract

The alarming rise in assaults against healthcare professionals is a public health and occupational issue that threatens staff well-being and care quality. Violence in this sector includes physical, verbal, and psychological aggression, posing a serious risk. Four main types of workplace violence in healthcare have been identified: External violence with no prior relationship (Type I), violence by patients against professionals (Type II, the most frequent), internal or institutional violence (Type III), and personal violence (Type IV). This issue is global, with an increasing trend and significant underreporting. Its consequences are severe at multiple levels: individually (burnout, anxiety, depression), institutionally (absenteeism, staff turnover), and in patient care quality. Artificial intelligence (AI) has emerged as a promising tool to prevent and mitigate such violence. Its applications include surveillance and monitoring systems, enhanced communication between staff and patients, workflow optimization, staff training, and predictive analysis of potentially aggressive patients. However, AI implementation presents ethical challenges related to data protection, privacy, bias risks, prediction reliability, and potential dehumanization. Addressing these concerns is crucial to ensuring safe and equitable AI use, always under human supervision. Effective prevention requires a comprehensive approach that integrates technology with organizational and educational measures.

KEYWORDS: Artificial intelligence. Aggressions. Health professionals. Prevention. Workplace violence. Underreporting.

*Correspondencia:

Juan Gómez-Salgado
E-mail: salgado@uhu.es

Fecha de recepción: 28-01-2025

Fecha de aceptación: 09-04-2025

DOI: 10.24875/GMM.M25001014

Gac Med Mex. 2025;161:573-581

Disponible en PubMed

www.gacetamedicademexico.com

0016-3813/© 2025 Academia Nacional de Medicina de México, A.C. Publicado por Permanyer. Este es un artículo open access bajo la licencia CC BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Introducción

Las agresiones al personal sanitario, ya sea manifestadas mediante violencia física o cualquier otro tipo de violencia, han aumentado de forma alarmante en las últimas décadas,¹ convirtiéndose en un importante problema de salud pública y laboral. Este fenómeno no solo afecta al bienestar del personal sanitario, sino que también compromete la calidad de la atención brindada a los pacientes. En este contexto, la inteligencia artificial (IA) surge como una herramienta prometedora y potencialmente transformadora para ayudar a prevenir y mitigar estos eventos, ofreciendo soluciones innovadoras para transformar la gestión de la seguridad en los entornos sanitarios.²

La violencia en el entorno laboral en el ámbito sanitario es un problema que va en aumento, con graves implicaciones para la salud física y mental de los profesionales afectados, la calidad asistencial y los sistemas de salud. Este fenómeno complejo, persistente y que prácticamente no se declara requiere intervenciones innovadoras que aprovechen las nuevas tecnologías, en particular la IA, para su identificación temprana, prevención y gestión eficaz.

Definición del fenómeno

La violencia en el entorno sanitario trasciende la mera agresión física para abarcar un espectro más amplio de conductas perjudiciales. Específicamente, la violencia contra los trabajadores sanitarios comprende cualquier acto de agresión –física, verbal o psicológica– capaz de manifestarse a través de distintos canales y con diferentes niveles de intensidad, dirigido contra profesionales sanitarios durante el ejercicio de sus funciones o como consecuencia de este.³⁻⁴ Este fenómeno se ha convertido en un asunto de creciente relevancia en el sector sanitario, reconocido como una amenaza notable para el bienestar de los profesionales y la calidad asistencial.⁵ Parte de la complejidad para estudiar este problema radica en la falta de una definición precisa que abarque la multiplicidad de formas en que se manifiesta.⁶ No obstante, existe consenso en considerar que estas agresiones constituyen una violación de derechos fundamentales con impacto tanto individual como colectivo, afectando no solo a la persona agredida, sino también al funcionamiento del sistema sanitario en su conjunto.⁷⁻⁸

Tipología

La clasificación establecida por la *Occupational Safety and Health Administration* (OSHA)⁹ distingue la violencia laboral en el ámbito sanitario en 4 tipos: Tipo I (Violencia externa), que implica a agresores sin relación previa con la institución, a menudo vinculados a delitos como robos; Tipo II (Violencia usuario-profesional), la más frecuente, que ocurre entre trabajadores sanitarios y pacientes o sus allegados por tensiones relacionadas con el servicio; Tipo III (Violencia interna o institucional), que incluye conflictos entre empleados o abuso jerárquico; y Tipo IV (Violencia personal), donde el agresor tiene una relación personal con un empleado pero sin conexión directa con el lugar de trabajo (Tabla 1).

Prevalencia y problema del que no se habla

El fenómeno de la violencia laboral en el ámbito sanitario ha alcanzado niveles alarmantes a nivel global, con una clara tendencia ascendente que pone en riesgo la sostenibilidad de los sistemas de salud.³⁻⁴ No obstante, un aspecto preocupante es la significativa infradeclaración de los incidentes, descrita a menudo en la literatura científica como “la punta del iceberg”, indicando que las estadísticas oficiales solo reflejan parte de la realidad.⁶⁻⁷

Los estudios muestran que un alto porcentaje de agresiones no se declara por diversas razones. Por ejemplo, un estudio en Andalucía encontró que el 67,6% de los incidentes no fueron registrados, ya que los trabajadores consideraban inútil participarlos o pensaban que la violencia era parte inherente del trabajo.⁸ Otros estudios en España e Italia también destacan tasas de infradeclaración de hasta el 73%.¹⁰⁻¹¹ Los factores que contribuyen a esta infradeclaración incluyen la insatisfacción con el manejo de incidentes previos, la ausencia de políticas que garanticen justicia, la percepción de que la violencia es normal y la dificultad o el tiempo requerido para completar el proceso de notificación.¹²

Perfiles y factores de riesgo

Los estudios muestran patrones definidos tanto en las víctimas como en los agresores. El perfil más frecuente del trabajador sanitario agredido corresponde a mujeres, con una media de edad de 37 años, predominantemente enfermeras que trabajan en

Tabla 1. Tipos de violencia en el lugar de trabajo

Tipo	Descripción	Ejemplo
I	El agresor no tiene ninguna relación con el lugar de trabajo ni con sus empleados.	Persona con intención delictiva comete un robo a mano armada.
II	El agresor es cliente o paciente del lugar de trabajo o de sus empleados.	Paciente intoxicado golpea a un auxiliar de enfermería.
III	El agresor es un empleado actual o anterior del lugar de trabajo.	Empleado recientemente despedido agrede a su antiguo supervisor.
IV	El agresor tiene una relación personal con un empleado, pero ninguna con el lugar de trabajo.	Exesposo agrede a su exesposa en su lugar de trabajo.

Adaptado de: European Agency for Safety and Health at Work.⁹

servicios de urgencias o atención primaria durante el turno de la mañana.¹

En cuanto a los agresores, el perfil predominante es el de pacientes o familiares, mayoritariamente hombres, con una media de edad de 42,5 años.³ Los servicios de urgencias y emergencias aparecen sistemáticamente como áreas de mayor riesgo, posiblemente debido a la combinación de estrés, tiempos de espera, gravedad de los casos atendidos y estados emocionales alterados.⁸

La violencia contra los profesionales de la salud es un problema global cuya manifestación se ve influida por contextos socioculturales y características del sistema sanitario. En Europa, los estudios indican que entre el 48 y el 65% de los trabajadores sanitarios sufren violencia laboral, siendo Italia y el Reino Unido países con tasas elevadas de incidentes.^{9,13-14} España registró un total de 14.749 casos en 2023, principalmente en atención primaria,^{1,15} impulsados por factores tales como la sobrecarga laboral y la insatisfacción de los pacientes.¹ En América Latina, países como Argentina y Perú informan tasas igualmente altas, hasta el punto de que el 89,56% del personal sanitario de un hospital público en Corrientes fue testigo de hechos violentos durante el último año¹⁶ y un 11,1% de trabajadores sanitarios peruanos llegaron a sufrir agresiones físicas.¹⁷ En Estados Unidos, el 81,6% de las enfermeras sufrieron episodios de violencia laboral en 2023.¹⁸

Por su parte, los trabajadores sanitarios de Asia y África enfrentan mayores riesgos debido a los conflictos armados que sufren y a lo debilitados que están sus

sistemas sanitarios, como evidencia el 70% del personal sanitario en Nigeria que participó hechos violentos en el trabajo.¹⁹ Estos hallazgos reflejan el alcance global del problema, con variaciones regionales. La implementación de políticas específicas para proteger a los profesionales de la salud es crucial para garantizar su seguridad y bienestar. Por otro lado, la tendencia temporal de los casos descritos indica un aumento en los episodios de violencia contra profesionales sanitarios, lo cual sugiere un deterioro progresivo en la relación entre los usuarios y los profesionales del sistema sanitario, posiblemente agravado por una mayor presión asistencial, recursos limitados y expectativas no satisfechas (Tabla 2).^{4,5,20-23}

Consecuencias

La violencia laboral en el ámbito sanitario tiene consecuencias de gran alcance e influye en los profesionales aumentando el riesgo de *burnout* (agotamiento emocional 37,6%, despersonalización 59,6%, baja realización personal 44,0%), ansiedad, depresión y una peor la calidad de vida.^{3,7} Las instituciones se enfrentan a más tasas de absentismo, un peor rendimiento laboral, deterioro del clima organizacional, una mayor rotación del personal y altos costes asociados a la gestión de incidentes.^{5,7} Además, la atención al paciente se ve afectada por una comunicación deteriorada entre profesional y paciente, actitudes defensivas y una menor confianza en la relación terapéutica.^{7,8}

Aplicaciones de la IA para la prevención de agresiones

La complejidad y magnitud de este problema requieren abordajes innovadores y multidimensionales. En este contexto, las tecnologías basadas en inteligencia artificial (IA) emergen como herramientas prometedoras para la identificación temprana de situaciones de riesgo, la prevención de incidentes y la gestión eficiente de eventos adversos en el entorno sanitario.^{2,22-28}

Las aplicaciones de la IA en este campo pueden clasificarse en cinco categorías principales (Tabla 3):

Sistemas avanzados de seguimiento y monitorización

Una de las principales formas en que la IA podría contribuir a la prevención de agresiones en entornos sanitarios es mediante la implementación de sistemas avanzados de seguimiento y monitorización.² Estas soluciones de seguridad impulsadas por IA pueden analizar imágenes de

Tabla 2. Datos seleccionados según contexto geográfico

Dato	Contexto geográfico	Fuente oficial	Fecha del dato
Estimaciones de prevalencia a un año en 61.800 profesionales sanitarios de 30 países. Alrededor de uno de cada cinco sufrió violencia física en el trabajo.	Global	Revisión sistemática y metaanálisis ²⁰	2000–2018
Hasta el 38% de los trabajadores sanitarios sufre violencia física en algún momento de su carrera.	Global	Organización Mundial de la Salud ²¹	2023
El sector salud y asistencia social registró las mayores tasas de violencia laboral entre todas las industrias privadas en EE. UU. durante el periodo 2021–2022.	Estados Unidos	Oficina de Estadísticas Laborales de EE. UU. ²²	2022
En 2018, los trabajadores sanitarios representaron el 73% de todas las lesiones y enfermedades laborales no fatales por violencia.	Estados Unidos	Oficina de Estad	
El 23% de los trabajadores de la salud de la UE "experimentaron al menos una forma de comportamiento social negativo durante los últimos 12 meses de trabajo".	Europa	European Nurses Research Foundation ¹²	2021
El 4,8% de los trabajadores de la salud activos sufre violencia verbal o física.	Europa	EU-OSHA ¹³	2020-2022
La violencia contra el personal sanitario ha aumentado, con un "23% de los trabajadores de la salud de la UE experimentando al menos una forma de comportamiento social negativo durante los últimos 12 meses de trabajo".	Europa	European Federation or Nurses Association ¹⁴	2024

Elaboración del propio autor basada en los datos referenciados.

Tabla 3. Usos potenciales de la inteligencia artificial (IA)

Sistemas avanzados de seguimiento y monitorización
Comunicación entre personal sanitario y pacientes
Optimización de flujos de trabajo
Formación y capacitación del personal sanitario
Análisis predictivo de pacientes potencialmente agresivos

Elaboración propia.

video en tiempo real para detectar comportamientos inusuales o potenciales amenazas, lo que permite la intervención inmediata del personal de seguridad.

Los algoritmos de IA pueden entrenarse para reconocer patrones de ansiedad, nerviosismo o agresividad en las interacciones entre el personal sanitario y los pacientes o sus acompañantes, lo cual posibilitaría alertas antes de la ocurrencia de posibles conflictos.^{2,24} Por ejemplo, el uso de técnicas de visión por computadora y aprendizaje profundo, como las redes neuronales convolucionales, facilita el análisis de transmisiones en vivo para detectar actividades inusuales y reducir la necesidad de supervisión humana continua.^{29,30}

La integración de la IA en los sistemas de vigilancia ofrece varios beneficios:

- Menor supervisión humana: al automatizar la detección de comportamientos anormales, se reduce la carga de trabajo del personal de seguridad, minimizando la fatiga y los errores humanos.^{29,31}
- Mayor precisión: los sistemas basados en IA pueden analizar grandes volúmenes de datos visuales con mayor precisión que los observadores humanos, lo cual disminuye las falsas alarmas y mejora la eficiencia en la detección de amenazas.³²
- Capacidad de aprendizaje: los algoritmos de IA pueden aprender de los datos recopilados y mejorar su capacidad para reconocer patrones de comportamiento inusual con el tiempo, permitiendo la adaptación continua a nuevos escenarios.³²

Aunque los sistemas de vigilancia avanzada ofrecen ventajas significativas, también presentan desafíos que tienen que ver con la privacidad y la ética, sobre todo, en entornos sanitarios donde la confidencialidad es primordial. Su efectividad depende de la calidad de los datos de entrenamiento y de su integración con los sistemas de seguridad existentes. No obstante, cuando se implementan adecuadamente, el seguimiento y monitorización potenciados por IA pueden llegar a ser herramientas poderosas para prevenir la violencia en entornos sanitarios, al permitir una detección temprana y precisa de comportamientos inusuales

Tabla 4. Estrategias basadas en IA para el manejo de la agresión en entornos psiquiátricos

Estrategia	Descripción	Ejemplo
Análisis predictivo	Predicción de riesgo basada en datos históricos y clínicos	Random forest aplicado a pacientes con esquizofrenia
Simulaciones virtuales	Formación interactiva para la desescalada	Escenarios con pacientes virtuales moderados por IA
Vigilancia avanzada	Detección temprana de conductas agresivas	Sistemas basados en visión artificial

Elaboración propia.

o amenazantes. Para maximizar sus beneficios y proteger la privacidad de los usuarios, es esencial abordar los desafíos éticos y técnicos asociados.^{29,33}

Mejora de la comunicación entre personal sanitario y pacientes

La implementación de tecnologías basadas en IA puede jugar un papel esencial en la mejora de la comunicación entre profesionales de la salud y pacientes. Esta comunicación efectiva es esencial para prevenir malentendidos que podrían desencadenar situaciones de conflicto. Las herramientas de procesamiento de lenguaje natural (PLN) pueden utilizarse para analizar interacciones con pacientes y sus cuidadores, identificando emociones que puedan indicar frustración o insatisfacción.³⁴ El uso del PLN permite analizar el lenguaje utilizado en las interacciones para detectar señales de malestar emocional. Al identificar estas señales en etapas tempranas, los profesionales sanitarios pueden abordar oportunamente la situación, reduciendo así la probabilidad de que surjan conflictos.²⁵ Por ejemplo, un estudio sobre el uso de modelos de lenguaje como GPT3.5 en la prevención de la diabetes demostró que las herramientas de PLN son capaces de dar respuestas personalizadas y apoyo emocional a los usuarios, lo cual sugiere la existencia de un potencial para mejorar la comunicación en otros contextos sanitarios.³⁵

La mejora en la comunicación no solo reduce los conflictos, sino que también contribuye a una mejor experiencia del paciente y a una mayor satisfacción con la atención recibida. Un estudio sistemático sobre los factores que influyen en la experiencia del paciente destacó la importancia de la comunicación profesional-paciente para mejorar la calidad de la atención sanitaria.³⁶ Asimismo, una comunicación

efectiva permite anticipar y resolver problemas antes de que escalen, lo que mejora el entorno laboral y reduce el estrés en el ámbito sanitario.

Por otro lado, es fundamental garantizar que estas herramientas sean accesibles para todos los usuarios, con independencia de su nivel de alfabetización digital o características culturales. Además, la implementación de estas tecnologías debe ir acompañada de políticas claras en materia de protección de datos y transparencia en el uso de la información recopilada. No obstante, es importante abordar los desafíos éticos y técnicos asociados para asegurar su implementación exitosa y el respeto a la privacidad del usuario.

Optimización de los flujos de trabajo

La optimización de los flujos de trabajo es un área clave en la que la IA puede contribuir de manera significativa al sector sanitario. Al automatizar tareas administrativas rutinarias, tales como la programación, gestión de documentación y otras tareas repetitivas o adaptables, la IA permite que los profesionales sanitarios estén más centrados en el cuidado del paciente y no en tarea administrativas.^{26,37} Los algoritmos de IA aplicados a la gestión sanitaria han demostrado efectos significativos en la reducción de situaciones potencialmente conflictivas.

La automatización de tareas administrativas no solo reduce la carga laboral de los profesionales de la salud, sino que también mejora la satisfacción laboral y reduce el riesgo de incidentes asociados al estrés.^{27,38} Al minimizar las cargas burocráticas, los profesionales pueden dedicar más tiempo a actividades clínicas, así como a una atención más directa, mejorando, así, la calidad del servicio y reduciendo el agotamiento emocional. Los sistemas basados en IA también optimizan la asignación de recursos al garantizar que los centros sanitarios cuenten con el personal y el equipamiento adecuados para satisfacer las necesidades de los pacientes. Esto incluye la gestión proactiva de expectativas, la priorización de casos según criterios clínicos y la distribución eficiente de recursos humanos y materiales.

Los estudios indican que la optimización del flujo de trabajo mediante IA puede reducir significativamente el estrés y el síndrome de *burnout* entre los profesionales de la salud. Por ejemplo, investigaciones sobre herramientas automatizadas de documentación en dermatología han demostrado que estas tecnologías alivian la carga administrativa y permiten a los médicos centrarse en la interacción con el paciente, lo cual reduce el síndrome de *burnout* y mejora la satisfacción laboral.³⁹ No obstante, la implementación exitosa requiere superar

desafíos tales como la integración con sistemas heredados, la protección robusta de la privacidad de los datos y programas de formación continua para maximizar los beneficios de la IA. Abordar estos desafíos resulta esencial para una adopción efectiva y sostenible de soluciones de IA en todo el ecosistema sanitario.

Formación y capacitación del personal sanitario

La formación y capacitación del personal sanitario son aspectos fundamentales para la prevención de agresiones en el sector salud. La IA desempeña un papel clave en este ámbito al mejorar la capacidad del personal para gestionar situaciones conflictivas. Las plataformas de simulación basadas en IA han revolucionado la formación del personal sanitario, permitiéndoles practicar y perfeccionar sus habilidades en entornos seguros y controlados.

Estas plataformas permiten a los profesionales de la salud ensayar escenarios altamente complejos, tales como el manejo de pacientes agitados o la resolución de conflictos interpersonales, sin riesgos para pacientes reales. Un estudio sobre el uso de simulación en la gestión de emergencias médicas en radiología halló que este tipo de formación mejoró significativamente los conocimientos y la confianza de los profesionales en el manejo de situaciones críticas, tales como paradas cardíacas y obstrucciones de la vía aérea.⁴⁰ Al dotar al personal sanitario de las habilidades necesarias para enfrentarse a situaciones difíciles, se fomenta un entorno laboral basado en la seguridad y el respeto mutuo, lo cual no solo reduce el riesgo de agresiones, sino que también mejora la satisfacción laboral y disminuye el estrés laboral. Un estudio sobre formación en resolución de conflictos en el ámbito hospitalario demostró que las intervenciones educativas podrían disminuir los conflictos relacionales, de tareas y procedimentales, fomentando estrategias de solución de problemas.⁴¹

Aunque las plataformas de simulación basadas en IA ofrecen beneficios sustanciales en materia de formación sanitaria, también presentan desafíos para su implementación, incluida una integración efectiva con los planes de estudio establecidos y la necesidad de una infraestructura tecnológica adecuada. Superar estos obstáculos es esencial para garantizar que estas tecnologías de formación avanzadas puedan implementarse de manera efectiva y sostenida en los programas educativos sanitarios.

Análisis predictivo de pacientes potencialmente agresivos

La aplicación de la IA en entornos sanitarios también puede contribuir al desarrollo de herramientas de análisis predictivo que identifiquen a pacientes predispuestos a conductas agresivas. Al analizar datos históricos y antecedentes clínicos, los modelos de IA pueden detectar patrones conductuales y pacientes con historial de comportamientos inusuales, lo cual permite al personal sanitario tomar medidas preventivas con antelación.

Estos modelos suelen basarse en técnicas de aprendizaje no supervisado, que permiten identificar patrones en grandes volúmenes de datos sin necesidad de etiquetado previo. Un estudio publicado en *JAMA Network Open* utilizó un abordaje de aprendizaje automático para evaluar el riesgo de violencia en pacientes hospitalizados a través de notas clínicas registradas rutinariamente en las historias clínicas electrónicas.²⁸ Este abordaje no solo mejora la seguridad ocupacional del personal sanitario, sino que también facilita la prestación de cuidados personalizados al asegurar que las personas reciban el nivel de atención y apoyo adecuado.⁴²

La identificación temprana de pacientes en riesgo de conducta agresiva permite adoptar medidas proactivas, tales como la asignación de personal adicional o la aplicación de protocolos específicos para gestionar situaciones potencialmente conflictivas. Este abordaje también contribuye a una atención más personalizada, ya que permite adaptar el nivel de cuidados a las necesidades de cada paciente. Un estudio sobre el uso del análisis predictivo en el manejo de pacientes críticos halló que las herramientas basadas en IA mejoran notablemente la precisión en la identificación de pacientes de alto riesgo, lo cual permite intervenciones tempranas y personalizadas.⁴³

La implementación de herramientas de análisis predictivo basadas en IA en entornos sanitarios presenta tanto oportunidades como desafíos. Aunque estas tecnologías mejoran la identificación de pacientes en riesgo de tener conductas agresivas, también generan preocupaciones importantes en materia de privacidad de datos y precisión de los modelos. Las organizaciones sanitarias deben priorizar la transparencia del sistema, la fiabilidad de las predicciones y la confidencialidad del paciente para mantener los estándares éticos durante su implementación. La integración exitosa de estas herramientas analíticas depende de superar problemas de compatibilidad con los flujos de trabajo y navegar por complejos

marcos regulatorios que actualmente limitan su adopción generalizada.^{42,43}

El desarrollo de sistemas predictivos potenciados por IA representa un abordaje estratégico para mejorar la seguridad ocupacional en entornos sanitarios. Al identificar a pacientes potencialmente agresivos antes de la ocurrencia de incidentes, estas herramientas permiten intervenciones proactivas y facilitan una atención personalizada. El sector sanitario deberá compensar la innovación tecnológica con consideraciones éticas, abordando las limitaciones técnicas y los desafíos de privacidad para garantizar que estos sistemas funcionen eficazmente, respetando los estándares de confidencialidad del paciente. El éxito depende, en última instancia, de crear marcos de implementación que avancen simultáneamente tanto en el manejo sanitaria como en la protección de los derechos de privacidad de cada paciente.^{44,45}

Estrategias basadas en IA para la gestión de agresiones en entornos psiquiátricos

Los entornos de salud mental plantean desafíos particulares e intensificados en relación con la violencia laboral, y la evidencia sugiere que los profesionales de salud mental corren hasta 3 veces más riesgo de agresión por parte de sus pacientes que en otros contextos sanitarios. Este riesgo elevado se atribuye a diversos factores, como las características de los trastornos psiquiátricos (impulsividad, paranoia o psicosis), el ingreso involuntario de muchos pacientes, el impacto de los fármacos sobre el comportamiento de los pacientes y la tensión inherente de las relaciones terapéuticas que tratan preocupaciones psicológicas complejas.⁴⁶⁻⁴⁸

En estos entornos, donde la agresión es una preocupación latente, la tecnología de IA ofrece herramientas especialmente prometedoras para mitigar estos riesgos, considerando las complejas conexiones entre factores clínicos, ambientales e interpersonales que suelen desencadenar incidentes violentos (Tabla 4). A diferencia de los métodos tradicionales de evaluación del riesgo, que se basan en información estática, las técnicas impulsadas por IA pueden combinar diversas fuentes de datos en tiempo real, tales como historiales médicos, observaciones de comportamiento actuales y tendencias previas para proporcionar valoraciones de riesgo más detalladas y oportunas. Estas aplicaciones son particularmente útiles en la atención psiquiátrica, donde los métodos de prevención convencionales pueden verse limitados por las características cognitivas y

conductuales de un grupo determinado de pacientes. Dichas aplicaciones especializadas de IA incluyen:

- Análisis predictivo del comportamiento agresivo: los modelos de aprendizaje automático, como random forest o máquinas de vectores de soporte, han demostrado una alta precisión a la hora de predecir conductas violentas en pacientes psiquiátricos, sobre todo, en aquellos con esquizofrenia o antecedentes de violencia. Por ejemplo, un estudio reciente mostró que el algoritmo random forest logró un área bajo la curva de 0,955 para predecir la ocurrencia de agresiones en entornos hospitalarios psiquiátricos.^{49,50}
- Plataformas de evaluación del riesgo: herramientas como *Electronic Hamilton Anatomy of Risk Management* integran datos clínicos e históricos para predecir y gestionar riesgos específicos de violencia en pacientes psiquiátricos, permitiendo intervenciones personalizadas y oportunas.⁵¹
- Simulaciones con pacientes virtuales: simulaciones impulsadas por IA que replican comportamientos agresivos típicos de pacientes psiquiátricos han demostrado ser eficaces para entrenar al personal sanitario en técnicas de desescalada y gestión emocional.
- Vigilancia continua mediante sistemas avanzados: sistemas de vigilancia basados en visión artificial permiten detectar signos tempranos de agitación o agresión, posibilitando intervenciones preventivas antes de la ocurrencia de incidentes violentos.^{50,52}

Limitaciones y consideraciones éticas

La integración de tecnologías basadas en IA para prevenir agresiones en el ámbito sanitario incorpora desafíos éticos cruciales incluidas preocupaciones sobre la privacidad de los datos, ya que los sistemas de videovigilancia y análisis conductual procesan información biométrica y de comportamiento sensible, lo cual plantea cuestiones de confidencialidad, tal y como demuestra un estudio chino sobre el uso de datos biomédicos con fines de seguridad nacional y corporativos.⁵³ Además, los algoritmos mal diseñados pueden reforzar sesgos preexistentes, siendo el sesgo y la discriminación problemas éticos ampliamente documentados en distintos sectores, incluido el sanitario.⁵⁴ El uso de análisis automatizado de expresiones faciales y comportamientos también plantea interrogantes sobre la autodeterminación cognitiva y la neuroética, en particular sobre la privacidad mental y la autonomía individual.⁵⁵ Por otro lado, las capacidades predictivas de la IA son limitadas y las

predicciones inexactas podrían influir en la autonomía o dignidad de los pacientes, razón por la cual la precisión y transparencia son esenciales.⁵⁶ Una dependencia excesiva de la tecnología también puede llevar a descuidar aspectos humanos y organizativos, lo cual puede llegar a deshumanizar la práctica médica y reducir la calidad de las interacciones profesional-paciente. Por último, abordar los sesgos algorítmicos es crucial para garantizar la equidad y prevenir decisiones erróneas o discriminatorias capaces de influir en poblaciones vulnerables. En este sentido, los sistemas de IA deben implementarse siempre bajo la supervisión de profesionales de la salud, asegurando que las decisiones finales estén guiadas por el juicio humano y la equidad.

Conclusiones

La IA tiene el potencial de reducir la sobrecarga profesional y prevenir agresiones mediante la predicción y mitigación de situaciones estresantes, mejorando así la seguridad asistencial, la calidad del cuidado y la eficiencia de los recursos. Cuando se incorpora a una estrategia integral con factores organizativos, formativos y ambientales, la IA puede jugar un papel clave en la prevención de la violencia contra los profesionales de la salud.

Algunas prometedoras aplicaciones incluyen la vigilancia inteligente, la optimización de flujos de trabajo, la mejora de la comunicación, la formación basada en simulación y los modelos predictivos de riesgo. No obstante, los desafíos éticos y técnicos deben ser abordados por vía multidisciplinar con profesionales sanitarios, expertos tecnológicos, especialistas en ética y representantes de pacientes.

En última instancia, la prevención efectiva de las agresiones requiere más que tecnología; exige una estrategia holística que incorpore dotación adecuada de personal, formación en comunicación, diseño de espacios seguros y políticas de tolerancia cero frente a la violencia.

Financiamiento

Ninguna.

Conflicto de intereses

Ninguno.

Consideraciones éticas

Protección de seres humanos y animales. Los autores declaran que no se realizaron experimentos

en seres humanos ni en animales para esta investigación.

Confidencialidad, consentimiento informado y aprobación ética. Este estudio no involucra datos personales de pacientes ni requiere aprobación ética. Las directrices SAGER no aplican.

Declaración sobre el uso de IA. Los autores declaran que no se utilizó IA generativa en la redacción de este manuscrito.

Referencias

1. Ministerio de Sanidad. Informe de agresiones a profesionales del Sistema Nacional de Salud 2023 [Internet]. Madrid: Ministerio de Sanidad; 2024 [citado el 2025 Mar 24]. Disponible en: <https://www.sanidad.gob.es/areas/profesionalesSanitarias/agresiones/home.htm>
2. Bhagat SV, Kanyal D. Navigating the Future: The Transformative Impact of Artificial Intelligence on Hospital Management- A Comprehensive Review. *Cureus*. 2024 Feb 20;16(2):e54518. doi: 10.7759/cureus.54518.
3. Yancán IA, Mayorca SL, García ME, Martínez ME, Lucerna Méndez MA. Violencia contra trabajadores sanitarios en un hospital de Andalucía: ¿Por qué hay agresiones no registradas? *Med Segur Trab*. 2013;59:235-258.
4. Barroso TC. Violencia laboral externa en el ámbito de urgencias generales del Hospital Universitario Virgen del Rocío de Sevilla. *Med Segur Trab*. 2016;62:25-48.
5. Silva-Bajaña JJ, Quinde-Alvear AG. Riesgos psicosociales y su impacto en el clima laboral del personal sanitario, en el área de emergencia del distrito de salud 08d05. *MQRInvestigar*. 2024;8(4):4581-4612. doi:10.56048/MQR20225.8.4.2024.4581-4612.
6. Cevallos Cárdenas KJ, Viteri Rojas AM, Valencia Cevallos AS, Rodríguez Jaramillo MDC. La violencia intrafamiliar. Definición y consecuencias. *RECIMUNDO* [Internet]. 2022 [citado 2025 Mar 24];6(4):384-97. Disponible en: <https://doi.org/10.26820/recimundo/6.4.octubre.2022.384-397>.
7. Vallejo-Andrada A, Sarasola Sánchez-Serrano JL, Barrera-Algarín E, Caravaca-Sánchez F. Burnout, ansiedad, satisfacción y violencia en el trabajo en trabajadores/as sociales de los servicios sociales municipales de Sevilla (España). *Alternativas Cuad Trab Soc*. 2025;32(1):81-105. doi:10.14198/ALTERN.25953.
8. Ceballos-Vásquez P, Espinoza-Arrué F, Solorza-Aburto J, Salgado-Vergara C, Almuna-Alfaro N, González-Andrade A, Muñoz-Rojas D. Violencia psicológica y física percibida por trabajadores de unidades de emergencias privadas de Chile. *Rev Cienc Salud*. 2024;22(2):1-16. <https://doi.org/10.12804/revistas.urosario.edu.co/revsalud/a.12528>.
9. European Agency for Safety and Health at Work. Workplace violence and harassment: a European picture [Internet]. 2011 [cited 2025 Mar 24]. Available from: <https://osha.europa.eu/en/publications/workplace-violence-and-harassment-european-picture>
10. Martínez Gandolfi MA, Rodríguez Mir J. Agresiones hacia el personal sanitario en España: violencias invisibilizadas, silencios inaceptables [Violence against healthcare workers in Spain: Invisible violence, unacceptable silences]. *Semergen*. 2024 Nov-Dec;50(8):102319. Spanish. doi: 10.1016/j.semgerg.2024.102319.
11. Veronesi G, Ferrario MM, Giusti EM, Borchini R, Cimmino L, Ghelli M, Banfi A, Luoni A, Persechino B, Di Tecco C, Ronchetti M, Gianfagna F, De Matteis S, Castelnuovo G, Iacoviello L. Systematic Violence Monitoring to Reduce Underreporting and to Better Inform Workplace Violence Prevention Among Health Care Workers: Before-and-After Prospective Study. *JMIR Public Health Surveill* 2023;9:e47377. doi: 10.2196/47377
12. Minha HS, Sahiran MN. Debunking the obstacles to reporting workplace violence among the healthcare workers in Melaka, Malaysia: a qualitative study. *Malaysian J Med Health Sci*. 2024 Mar 15;20(2):142-50. doi: 10.47836/mjmhs.20.2.19.
13. European Nurses Research Foundation. Workplace Violence Against Nurses in the European Union. 2022. Available from: <https://www.enrf.eu/wp-content/uploads/2022/11/ENRF-Policy-Brief-Violence-Against-Nurses-November-2022.pdf>
14. European Federation of Nurses Associations. European Federation of Nurses Associations [Internet]. [cited 2025 Mar 24]. Available from: <https://efn.eu>
15. Gascón-Santos S, Oliván-Blázquez B, Chambel M, Sebastián L, Alacreu-Crespo A, Pérez-Montesinos Y, et al. Symptomatology of post-traumatic stress and burnout associated with aggressions suffered by primary care health professionals. A transversal study. *Front Psychol*. 2024;15. doi:10.3389/fpsyg.2024.1335155
16. Bravo ED, Baez AN, Cozzoli BC, Auchter MC, Meza AM. Situaciones de violencia por parte de pacientes percibidos por personal de enfermería en un hospital público Corrientes capital 2021. *Notas Enferm (Córdoba)*. 2022 Dic;23(40):37-46.

17. Tuya-Figueroa X, Mezones-Holguin E, Monge E, Arones R, Mier M, Saravia M, Torres J, Mayta-Tristán P. Violencia laboral externa tipo amenaza contra médicos en servicios hospitalarios de Lima Metropolitana, Perú 2014. *Rev Peru Med Exp Salud Publica*. 2016 Dec 13;33(4):670. doi: 10.17843/rpmesp.2016.334.2550.
18. National Nurses United. Workplace Violence Report [Internet]. Oakland, California, USA: NNU; 2024 [cited Mar 24. Available from: https://www.nationalnursesunited.org/sites/default/files/nnu/documents/0224_Workplace_Violence_Report.pdf.
19. Elom P, Agu A, Unah A, Azuogu B, Ituma B, Okah O, Okocha Y, Ugwuweze J, Ossai E, Igwe D. Prevalence and factors associated with workplace violence in a tertiary healthcare facility in Nigeria. *Niger Med J*. 2024 May 23;65(2):173-184. doi: 10.60787/nmj-v65i2-336.
20. Li YL, Li RQ, Qiu D, Xiao SY. Prevalence of Workplace Physical Violence against Health Care Professionals by Patients and Visitors: A Systematic Review and Meta-Analysis. *Int J Environ Res Public Health*. 2020 Jan 1;17(1):299. doi: 10.3390/ijerph17010299.
21. World Health Organization. Preventing Violence Against Health Workers. Geneva: WHO; 2022. [Internet]. [cited 2025 Mar 24. Available from: <https://www.who.int/activities/preventing-violence-against-health-workers>
22. U.S. Bureau of Labor Statistics. Workplace Violence 2021-2022. [Internet]. Washington, USA: U.S. Bureau of Labor Statistics; 2022. Available from: <https://www.bls.gov/iif/factsheets/workplace-violence-2021-2022.htm>
23. U.S. Bureau of Labor Statistics. Workplace Violence in Healthcare, 2018. [Internet]. Washington, USA: U.S. Bureau of Labor Statistics; 2018. Available from: <https://www.bls.gov/iif/factsheets/workplace-violence-healthcare-2018.htm>
24. Bekbolatova M, Mayer J, Ong CW, Toma M. Transformative Potential of AI in Healthcare: Definitions, Applications, and Navigating the Ethical Landscape and Public Perspectives. *Healthcare (Basel)*. 2024 Jan 5;12(2):125. doi: 10.3390/healthcare12020125.
25. Chen M, Decary M. Artificial intelligence in healthcare: An essential guide for health leaders. *Health Manage Forum*. 2020 Jan;33(1):10-18. doi: 10.1177/0840470419873123.
26. Alowais SA, Alghamdi SS, Alsuhbany N, Alqahtani T, Alshaya AI, Almohareb SN, Aldairem A, Alrashid M, Bin Saleh K, Badreldin HA, Al Yamani MS, Al Harbi S, Albekairy AM. Revolutionizing healthcare: the role of artificial intelligence in clinical practice. *BMC Med Educ*. 2023 Sep 22;23(1):689. doi: 10.1186/s12909-023-04698-z.
27. Meskó B, Hetényi G, Györfy Z. Will artificial intelligence solve the human resource crisis in healthcare? *BMC Health Serv Res*. 2018 Jul 13;18(1):545. doi: 10.1186/s12913-018-3359-4.
28. Menger V, Spruijt M, van Est R, Nap E, Scheepers F. Machine Learning Approach to Inpatient Violence Risk Assessment Using Routinely Collected Clinical Notes in Electronic Health Records. *JAMA Netw Open*. 2019 Jul 3;2(7):e196709. doi: 10.1001/jamanetworkopen.2019.6709.
29. Narkhede VP, Jawandhiya PM. A systematic review on evaluation of surveillance system for unusual behavior using artificial intelligence. *Int J Recent Innov Trends Comput Commun*. 2023;11(11):1508-1513. doi: 10.17762/ijritcc.v11i11.10951.
30. Mohammad R, Saeed F, Almazroi AA, Alsubaie FS, Almazroi AA. Enhancing intrusion detection systems using a deep learning and data augmentation approach. *Systems*. 2024;12(3):79. doi: 10.3390/systems12030079.
31. Jain S, Choudhary N. AI techniques for anomaly detection in video surveillance using deep learning method. In: *Proceedings of the 2024 3rd International Conference on Artificial Intelligence for Internet of Things (AIoT)*; 2024 Mar; Vellore, India. IEEE; 2024. p. 1-6. doi: 10.1109/AlloT58432.2024.10574643.
32. Sivakumar C, Vali TK, Reddy PSB, Meghana ML, Sukumar Y. AI-Powered Video Surveillance for Enhanced Intrusion Detection. In: *Proceedings of the 2024 International Conference on IoT Based Control Networks and Intelligent Systems (ICINIS)*; 2024 Mar; Bengaluru, India. p. 1630-1634. doi: 10.1109/ICINIS64247.2024.10823272.
33. Ahmad I, Sahoo SR, Dash R. Anomaly detection in video surveillance for unusual behavior identification. In: *2024 2nd World Conference on Communication & Computing (WCONF)*; 2024 Mar 20-22; Raipur, India. New York: IEEE; 2024. p. 1-6. doi: 10.1109/WCONF61366.2024.10692100.
34. Alosban N, Esposito A, Vinciarelli A. What You Say or How You Say It? Depression Detection Through Joint Modeling of Linguistic and Acoustic Aspects of Speech. *Cogn Comput*. 2022;14(6):1585-1598. doi:10.1007/s12559-020-09808-3.
35. Dao D, Teo JYC, Wang W, Nguyen HD. LLM-Powered Multimodal AI Conversations for Diabetes Prevention. In: *Proceedings of the 1st ACM Workshop on AI-Powered Q&A Systems for Multimedia (AIQAM '24)*; 2024; New York, NY, USA. New York: Association for Computing Machinery; 2024. p. 1-6. doi: 10.1145/3643479.3662049.
36. Kim EJ, Koo YR, Nam IC. Patients and healthcare providers' perspectives on patient experience factors and a model of patient-centered care communication: a systematic review. *Healthcare*. 2024;12(11):1090. doi:10.3390/healthcare12111090.
37. Ko JS, Heo H, Suh CH, Yi J, Shim WH. Adherence of Studies on Large Language Models for Medical Applications Published in Leading Medical Journals According to the MI-CLEAR-LLM Checklist. *Korean J Radiol*. 2025 Jan 23. doi: 10.3348/kjr.2024.1161.
38. Ahmed H, Akber N, Saleem M, Ahmed F, Yasmeen R, Ali L. The Application of AI in Clinical Nursing, Yields Several Advantageous Outcomes. *Indus J Biosci Res*. 2025;3(2):591-9. doi:10.70749/ijbr.v3i2.731.
39. Roth J, Thunga S, Yoo J. Assessing the Landscape of AI-Powered Patient Documentation in Dermatology. *J Drugs Dermatol*. 2025 Jan 1;24(1):64-69. doi: 10.36849/JDD.8583.
40. Wheeler M, Powell E, Pallmann P PhD. Use of High-fidelity simulation training for radiology healthcare professionals in the management of acute medical emergencies. *Br J Radiol*. 2021 Jan 1;94(1117):20200520. doi: 10.1259/bjr.20200520.
41. Imani B, Sabet Ghadam M, Gomrokchi AY, Maddineshat M. Efficacy of tailored educational interventions in changing conflict management strategies among operating room staff. *Work*. 2024 Nov 29;10519815 241297844. doi: 10.1177/10519815241297844.
42. Dobbins NJ, Chipkin J, Byrne T, Ghabra O, Siar J, Sauder M, Huijton RM, Black TM. Deep learning models can predict violence and threats against healthcare providers using clinical notes. *Npj Ment Health Res*. 2024 Dec 5;3(1):61. doi: 10.1038/s44184-024-00105-7.
43. Janjua JI, Ghazal TM, Abushiba W, Abbas S. Optimizing patient outcomes with AI and predictive analytics in healthcare. 2024 IEEE 65th International Scientific Conference on Power and Electrical Engineering of Riga Technical University (RTUCon); 2024 Oct 9-11; Riga, Latvia. p. 1-6. doi: 10.1109/RTUCon62997.2024.10830874.
44. Gerke S, Babic B, Evgeniou T, Cohen IG. The need for a system view to regulate artificial intelligence/machine learning-based software as medical device. *NPJ Digit Med*. 2020 Apr 7;3:53. doi: 10.1038/s41746-020-0262-2.
45. Karimian G, Petelos E, Evers SMAA. The ethical issues of the application of artificial intelligence in healthcare: a systematic scoping review. *AI Ethics*. 2022;2:539-51. doi: 10.1007/s43681-021-00131-7.
46. Pekurinen V, Willman L, Virtanen M, Kivimäki M, Vahtera J, Välimäki M. Patient Aggression and the Wellbeing of Nurses: A Cross-Sectional Survey Study in Psychiatric and Non-Psychiatric Settings. *Int J Environ Res Public Health*. 2017 Oct 18;14(10):1245. doi: 10.3390/ijerph14101245.
47. Ridenour M, Lanza M, Hendricks S, Hartley D, Rierdan J, Zeiss R, Amandus H. Incidence and risk factors of workplace violence on psychiatric staff. *Work*. 2015;51(1):19-28. doi: 10.3233/WOR-141894.
48. Weltens I, Bak M, Verhagen S, Vandenberk E, Domen P, van Amelsvoort T, Drukker M. Aggression on the psychiatric ward: Prevalence and risk factors. A systematic review of the literature. *PLoS One*. 2021 Oct 8;16(10):e0258346. doi: 10.1371/journal.pone.0258346.
49. Cheng N, Guo M, Yan F, Guo Z, Meng J, Ning K, Zhang Y, Duan Z, Han Y, Wang C. Application of machine learning in predicting aggressive behaviors from hospitalized patients with schizophrenia. *Front Psychiatry*. 2023 Mar 20;14:1016586. doi: 10.3389/fpsyt.2023.1016586.
50. Tay JL, Li Z, Sim K. Effectiveness of Artificial Intelligence Methods in Personalized Aggression Risk Prediction within Inpatient Psychiatric Treatment Settings-A Systematic Review. *J Pers Med*. 2022 Sep 7;12(9):1470. doi: 10.3390/jpm12091470.
51. Chaimowitz G, Mamak M, Moulden H. Using artificial intelligence to improve prediction and prevention of violence. *Res Outreach*. 2021;(124):1-4. Available from: <https://researchoutreach.org/articles/using-artificial-intelligence-improve-prediction-prevention-violence>
52. Idemudia ES. Journal abstracts from current research in the field of child and adolescent mental health. *J Child Adolesc Ment Health*. 2013;25(1):99-103. doi: 10.2989/17280583.2013.802443.
53. Char D. Challenges of Local Ethics Review in a Global Healthcare AI Market. *Am J Bioeth*. 2022 May;22(5):39-41. doi: 10.1080/15265161.2022.2055214.
54. Burema D, Debowski-Weimann N, von Janowski A, Grabowski J, Maftai M, Jacobs M, van der Smagt P, Benbouzid D. A sector-based approach to AI ethics: Understanding ethical issues of AI-related incidents within their sectoral context. In: *Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society (AIES '23)*; 2023 Mar 7-10; New York, NY, USA. New York (NY): Association for Computing Machinery; 2023. p. 705-14. doi: 10.1145/3600211.3604680.
55. Kumar A, Braud T, Tarkoma S, Hui P. Trustworthy AI in the Age of Pervasive Computing and Big Data. En: 2020 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops); 2020 Mar 23-27; Austin, TX, USA. IEEE; 2020. p. 1-6. doi: 10.1109/PerComWorkshops48775.2020.9156127.
56. Blasimme A, Vayena E. The ethics of AI in biomedical research, patient care, and public health. In: Dubber MD, Pasquale F, Das S, eds. *The Oxford Handbook of Ethics of AI*. 2020. Oxford Academic; 9 Jul 2020. doi:10.1093/oxfordhb/9780190067397.013.45