

Revisión sistemática sobre el uso de machine learning (ML) para la predicción de conductas autolesivas y suicidas en adolescentes

José Manuel Granada López

Primer premio de Investigación Enfermera

"D. Ángel Andía Leza" 2025

del Colegio Oficial de Enfermería de Zaragoza

RESUMEN

Introducción: La conducta suicida y las autolesiones representan un grave problema de salud pública en adolescentes, con una prevalencia creciente en los últimos años. Esta revisión sistemática tuvo como objetivo analizar la efectividad de los modelos de aprendizaje automático (machine learning, ML) para predecir dichas conductas en población adolescente. Se identificaron 24 estudios publicados entre 2021 y 2025, seleccionados tras una búsqueda en las bases de datos PubMed, Scopus, Web of Science y PsycINFO.

Los estudios incluidos emplearon diversos diseños metodológicos, destacando los observacionales retrospectivos y longitudinales, y utilizaron una amplia variedad de algoritmos de ML como random forest, redes neuronales profundas (DNN), support vector machines (SVM) y XGBoost. Las métricas de desempeño reportadas fueron principalmente AUC, sensibilidad y precisión, con valores que oscilaron entre 0.70 y 0.93.

Las variables predictoras más frecuentes fueron síntomas depresivos, antecedentes de autolesión, bajo apoyo social y diagnóstico psiquiátrico previo. Aunque los resultados muestran una capacidad predictiva prometedora, las limitaciones comunes incluyen el uso de autoinformes, el sesgo de selección y la escasa validación externa.

En conjunto, los modelos de ML representan una herramienta emergente con potencial para complementar las estrategias tradicionales de prevención, especialmente si se abordan los retos éticos y metodológicos pendientes. La integración de estos modelos en contextos clínicos y comunitarios podría contribuir a la detección temprana y a la intervención personalizada en adolescentes en riesgo.

Palabras clave: Suicidio, autolesiones, adolescentes, machine learning, inteligencia artificial, predicción, salud mental, prevención.

■ INTRODUCCIÓN

1.1. Conducta suicida y autolesiones en adolescentes

El suicidio, como conducta que expresa el pensamiento que busca la propia muerte, es una de las principales causas de mortalidad en la adolescencia a nivel mundial. Así, la Organización Mundial de la Salud (OMS), en su informe "Health for the World's Adolescence", pone el foco, entre otras causas de muerte adolescente, en el suicidio (World Health Organization, 2014).

Liu et al. en su metaanálisis de 2022 sobre 58 estudios en menores de 13 años, inicio de la adolescencia, la prevalencia de ideación suicida es del 15,1%, las tentativas suicidas constituyen un 2,6% y la conducta autolesiva con intención no suicida el 6,2%. Aunque las muertes por suicidio en esta población infrecuentes (0,79 por cada millón de niños), la prevalencia de autolesiones es preocupante. Los factores de riesgo identificados corresponden a sexo masculino, antecedentes de maltrato infantil, presencia de trastornos

como déficit de atención con hiperactividad (TDAH) y depresión, mientras que el apoyo de las figuras paternas destaca como un factor protector significativo (Liu et al., 2022).

El documento del CDC (Preventing Suicide: A Technical Package of Policies, Programs, and Practices, 2017) habla del suicidio en adolescentes, destacando que el suicidio es la tercera causa principal de muerte en jóvenes de 10 a 14 años en EE.UU. (Stone et al., 2017).

En España, Vázquez López et al. alertan sobre el aumento de las autolesiones y conductas suicidas en niños y adolescentes tras la pandemia, destacando la necesidad de diferenciar entre autolesiones con intención no suicida y conducta suicida, así como de reforzar la formación de pediatras y docentes. Señalan factores de riesgo como la depresión, tentativas previas y el acoso y abuso infantil, y reclaman un plan nacional de prevención del suicidio que integre recursos sanitarios, educativos y sociales (Vázquez López et al., 2023).

Por tanto, es éste un fenómeno complejo, con diversos determinantes asociados: patología mental, trauma derivado

de situaciones como exclusión, marginación, violencia, migración, abuso, etc. (Su et al., 2023; Smith et al., 2023). Las autolesiones, conductas de daño autoinfligido deliberadamente sin intención suicida, como alivio de tensión o expresión de malestar emocional, también han aumentado su prevalencia y, aunque no se puede hacer demasiados paralelismos con el suicidio, en ocasiones pueden ser predictores de la conducta suicida y se manifiestan fundamentalmente durante la adolescencia (Su et al., 2023; Guzmán Brand y Gélvez García, 2024).

1.2. Abordaje basado en la tecnología con métodos predictivos

Las acciones preventivas de suicidio implementadas hasta ahora han sido ineficaces debido a su diseño deficiente, la falta de una definición de caso uniforme y un enfoque reduccionista centrado únicamente en la salud pública y en los esquemas tradicionales de prevención primaria, secundaria y terciaria. Se sugiere ampliar la perspectiva para incluir aspectos filosóficos, psicosociales y psiquiátricos, e implementar estrategias preventivas basadas en intervenciones poblacionales que aborden de manera integral los diversos factores que contribuyen al comportamiento suicida (Campillo Serrano y Fajardo Dolci, 2021).

La publicación del CDC (Preventing Suicide: A Technical Package of Policies, Programs, and Practices, 2017) insiste en una aproximación multiescalar (individual, comunitaria y social) y multisectorial para abordar los factores de riesgo y de protección y enseñar habilidades de afrontamiento desde edades tempranas como estrategias clave de prevención (Stone et al., 2017).

El uso de estas nuevas tecnologías digitales (apps, plataformas online) tienen ya cierto recorrido en su implementación para la reducción de ideación suicida, con moderados resultados, es un área con gran potencial, aunque falta investigación más sólida (Gaynor et al., 2023) parece que va a ser el desarrollo de machine learning (ML), todavía en formas muy iniciales, la funcionalidad que permita identificar patrones predictivos, estudiar interacciones sociales y detectar textos o actividades que puedan ser indicadores de riesgo y basar en estos aplicativos la prevención de la conducta suicida (Guzmán Brand y Gélvez García, 2024; Lee et al., 2024; Zang et al., 2024).

1.3. Machine Learning

En los últimos años ha habido un cambio de paradigma en el estudio epidemiológico aplicado a la salud pública mediante el tratamiento avanzado de datos, el Machine Learning (ML), una parte de los desarrollos de la inteligencia artificial (IA), permite el análisis de gran cantidad de datos, de fuentes diversas, con características diferentes, y a una velocidad

altísima. Esto permite el estudio de correlaciones entre variables complejas que no podrían ser abordados mediante la metodología estadística tradicional (Servi et al., 2023; Kim et al., 2024; Mora Pineda, 2022). Estas funcionalidades son de gran interés en el ámbito de la salud mental, en el que la multicausalidad, la escasa linealidad entre relaciones de factores, ha dificultado los métodos predictivos tradicionales. Los modelos basados en ML, permiten la entrada de datos que ya no tienen que ir obligatoriamente en formato tabular, sino que permite trabajar con textos, imágenes, videos, series temporales, ficheros de audio, y tratamientos complejos de los mismos (Mora Pineda, 2022). Esta metodología se ha mostrado útil en la predicción de tentativas suicidas mediante el análisis de datos longitudinales, tomando diversos determinantes psicosociales, datos clínicos y de conductas (Zang et al., 2024; Yang et al., 2024).

Su et al. (2020) demostraron que, a partir de registros electrónicos de salud, es posible construir modelos predictivos de alta precisión para identificar tanto riesgo inmediato de conducta suicida como a mayor plazo. Destacan la utilidad del aprendizaje automático para detectar factores de riesgo relevantes, como manifestaciones de malestar emocional y episodios depresivos, facilitando las intervenciones tempranas y eficaces (Su et al., 2020).

1.4. Justificación

Las conductas suicidas y autolesivas han sido catalogadas como asuntos de salud pública, y la OMS insta a las autoridades sanitarias a intervenir, principalmente en acciones preventivas, de manera urgente (World Health Organization, 2019).

El reciente trabajo de Su et al. usa algoritmos de ML para la predicción de conductas autolesivas y suicidas en la adolescencia, tomando factores como la dinámica familiar, el entorno escolar y la sintomatología depresiva como indicadores predictores clave (Su et al., 2023). También los modelos basados en algoritmos y redes neuronales han sido considerados útiles en el estudio de diversas cohortes internacionales (Lee et al., 2024; Kim et al., 2024). Incluso, permite diversificar hasta tal punto las correlaciones que podemos leer recientes investigaciones como las de Lee et al., que utilizan el ML para analizar comorbilidades, como la rinitis alérgica, valorando su posible impacto en las ideaciones suicidas en la adolescencia (Lee et al., 2024).

También son novedosos los diseños en predicción de suicidio a partir de datos clínicos estructurados y no estructurados, como el uso de parámetros biológicos aplicados a la psicopatología, como Cansel et al. que analizan datos plasmáticos en sangre completa para la predicción suicida mediante IA, demostrando que hay biomarcadores de utilidad en la detección

temprana (Cansel et al., 2023). El trabajo de De Luca et al., en cambio, utiliza la herramienta Suicide Crisis Inventory-2 (SCI-2), basada en ML y su aplicación a la predicción de ideación suicida en la edad adolescente (De Luca et al., 2024).

Estas aplicaciones iniciales se topan con dificultades diversas, lo complejo del pensamiento y conducta de las personas, la cantidad y variabilidad de los datos y de sus fuentes, los sesgos a la hora de su tratamiento e interpretación, dificultarán la transferibilidad de estos estudios a la práctica clínica y preventiva habitual. (Smith et al., 2023; Haroz et al., 2024). Además, estamos ante el tratamiento de los datos de mayor vulnerabilidad y a los que se otorga la máxima protección, así que cuestiones éticas y legales de estos datos y todo lo relacionado con el estigma que puede producir el señalamiento de adolescentes identificados como de alto riesgo, también pueden dificultar enormemente la implementación de estos sistemas basados en ML (Smith et al., 2023; Kim et al., 2024).

Pero la aplicación de ML mejora los tratamientos tradicionales de datos clínicos, Zang et al. han usado registros clínicos longitudinales de forma útil para cohortes de distintos contextos socioculturales (Zang et al., 2024), y otros se han centrado en el estudio de redes neuronales con algoritmos de aprendizaje profundo, con buenos resultados en la predicción de tentativas suicidas adolescentes (Bazrafshan y Sayehmiri, 2024). Estos estudios van mostrando que el camino podría ser el uso de ML en la adaptación a cada sujeto de estrategias para la prevención y el establecimiento de indicadores segregados por género y edad, y otros determinantes, que permitan alertar de factores de riesgo (Kusuma et al., 2024).

Lu et al. han estudiado datos cuyas fuentes son las redes sociales y entornos sociales de los jóvenes y qué puede identificar factores de riesgo asociados a los pensamientos y conducta suicida (Lu et al., 2024). Y, si se va a identificar factores protectores, Haghish et al. destacan en su estudio la importancia de las dinámicas en el grupo de pertenencia y los apoyos y relaciones como freno a que la ideación suicida pueda pasar al acto (Haghish et al., 2024a).

Pero aún es difícil, como se ha visto, llevar a la práctica habitual estos modelos, los estudios disponibles son limitados, con baja representatividad de poblaciones y falta de estandarización de indicadores (Czyz et al., 2023). Aunque sí que se avanza en la aplicación de ML a datos que provengan de escalas, test y entrevista clínica, orientada a la predicción del riesgo suicida en un entorno clínico (Wilimitis et al., 2022; Jankowsky et al., 2024).

Puede ser una herramienta complementaria para prevenir una de las principales causas de mortalidad en una etapa vital especialmente sensible. El uso de ML puede ser importante en los contextos clínicos asistenciales, pero también en

prevención en población general, permitiendo que se llegue a casos que pasarían desapercibidos a las prácticas actuales habituales.

Es necesario perfeccionar el tratamiento de datos, los modelos de ML, los algoritmos, y las premisas éticas y legales del uso de datos personales, pero el potencial de esta tecnología es tal que justifica el análisis de la información disponible y el trazar líneas investigadoras futuras, según la OMS los estudios deben avalar la efectividad de las intervenciones preventivas, que es el grado en que logran los resultados deseados: reducir la incidencia, prevalencia o gravedad de una enfermedad, cuando se aplica en condiciones reales, no controladas, en poblaciones específicas (World Health Organization, 2016). En concreto, la efectividad de una intervención en prevención del suicidio se mide por su capacidad para reducir la ideación suicida, las tentativas o los suicidios consumados en una población real, comparado con no intervenir o con otra intervención estándar (Zalsman et al., 2016).

Además, engarza con el Objetivo 3: Garantizar una vida sana y promover el bienestar para todos en todas las edades, de los Objetivos de Desarrollo Sostenible de las Naciones Unidas.

Por tanto, conviene realizar una revisión sistemática de la bibliografía disponible actualmente sobre el uso de ML para la predicción de conductas autolesivas y suicidas en población adolescente.

■ OBJETIVOS

2.1 Objetivo general

- Evaluar la efectividad de los modelos de ML para la predicción de conductas autolesivas y suicidas en adolescentes.

2.2 Objetivos específicos

- Identificar los modelos de ML utilizados en estudios actuales sobre autolesiones y suicidios en adolescentes.
- Determinar indicadores clave asociados a factores de riesgo suicida.
- *Analizar las posibilidades de implementar estos modelos en diferentes contextos socioculturales.*
- Evaluar limitaciones y posibilidades de mejora en el uso de ML en la salud mental.

2.3 Preguntas de investigación

Para redactar de forma clara y precisa las preguntas de investigación, se empleó la metodología PICO, que facilita la definición de la población diana, la intervención empleada, si se puede comparar con otras alternativas y los resultados esperados.

Paciente o problema	Intervención (tratamiento, factor, causa, etc.)	Comparación (si procede)	Resultados (Outcome)
Adolescentes en riesgo de conductas autolesivas y suicidas.	Modelos de predicción basados en machine learning (ML) o inteligencia artificial (IA)	Métodos tradicionales de predicción (por ejemplo, juicio clínico o escalas de riesgo convencionales)	Efectividad en la predicción de conductas autolesivas y suicidas

¿Los modelos basados en ML o IA son más efectivos que los métodos tradicionales para predecir conductas autolesivas y suicidas en adolescentes?

Paciente o problema	Intervención (tratamiento, factor, causa, etc.)	Comparación (si procede)	Resultados (Outcome)
Adolescentes en riesgo de suicidio	Identificación de factores predictivos clave mediante machine learning (ML)	Identificación de factores mediante metodologías tradicionales (análisis estadístico convencional, juicio clínico)	Mejora en la efectividad de las estrategias de prevención del suicidio

¿La identificación de factores predictivos clave mediante ML mejora la efectividad de las estrategias de prevención del suicidio en adolescentes en comparación con metodologías tradicionales?

METODOLOGÍA

Se trata de un trabajo dirigido a la revisión sistemática de la bibliografía sobre un asunto concreto, aunque, dado que la realiza una sola persona y aporta evidencia media, se debería de hablar de una Revisión Sistematizada de la Literatura o de la bibliografía (RSL) (Masot y Selva-Pareja, 2020).

3.1. Estrategia de búsqueda

La revisión sistemática se realizó mediante la búsqueda de artículos, durante el mes de marzo y primera mitad de abril de 2025, en las siguientes bases de datos: PubMed, Scopus, Web of Science y PsycINFO. Se emplearon descriptores DeCS y MeSH relacionados con el tema de estudio, combinados con operadores booleanos para optimizar la recuperación de información relevante.

Las palabras clave incluyeron: Machine learning, artificial intelligence, deep learning, predictive modeling, self-harm, self-injurious behavior, suicide, teenagers, adolescents, youth, young people.

Operadores booleanos y metodología de búsqueda: AND para combinación de términos diferentes y OR para alternativas a términos similares.

Se aplicaron los siguientes filtros:

- Estudios en personas.
- Publicados en los últimos 5 años (esta tecnología tiene una evolución muy rápida por lo que no se han considerado publicaciones anteriores).
- Idiomas: inglés y español.
- Tipos de estudio: revisiones sistemáticas, ensayos clínicos, estudios longitudinales y casos de cohortes.
- Texto completo disponible (Free full text).
- Truncamiento con "Adolescent*" para que use "Adolescent" o "Adolescents".
- Entrecomillado para los términos de más de una palabra: "Machine Learning" o "Self-Injurious", por ejemplo.

Se resume la búsqueda y artículos encontrados en la Tabla 1.

Tabla 1. Estrategia de búsqueda bibliográfica.

Base de datos	Palabras clave	Filtros aplicados	Sintaxis utilizada	Artículos
PubMed	Machine Learning, Artificial Intelligence, Deep Learning, Predictive Modeling, Self-Harm, Self- Injurious Behavior, Suicide, Teenagers, Adolescents, Youth, Young People	Últimos 5 años, inglés/español, abstract disponible	("Machine Learning" [Mesh] OR "Artificial Intelligence" [Mesh] OR "Deep Learning" OR "Predictive Modeling") AND ("Suicide" [Mesh] OR "Self-Injurious Behavior" [Mesh]) AND ("Adolescent" [Mesh])	55
Scopus	Machine Learning, Artificial Intelligence, Deep Learning, Predictive Modeling, Self-Harm, Self- Injurious Behavior, Suicide, Teenagers, Adolescents, Youth, Young People	Artículos, revisiones y estudios originales. Medicina, Psicología, Ciencias Sociales. Últimos 5 años	TITLE-ABS-KEY ("machine learning" OR "artificial intelligence" OR "deep learning" OR "predictive modeling") AND TITLE-ABS-KEY ("suicide" OR "self-harm" OR "self-injurious behavior") AND TITLE- ABS-KEY ("adolescent*" OR "teenager*" OR "youth" OR "young people")	131
Web of Science	Machine Learning, Artificial Intelligence, Deep Learning, Predictive Modeling, Self-Harm, Self- Injurious Behavior, Suicide, Teenagers, Adolescents, Youth, Young People	Últimos 5 años, artículos, revisiones	("machine learning" OR "artificial intelligence" OR "deep learning" OR "predictive modeling") AND ("suicide" OR "self-harm" OR "self- injurious behavior") AND ("adolescent*" OR "teenager*" OR "youth" OR "young people")	14
PsycINFO	Machine Learning, Artificial Intelligence, Deep Learning, Predictive Modeling, Self-Harm, Self- Injurious Behavior, Suicide, Teenagers, Adolescents, Youth, Young People	Adolescentes, estudios empíricos	("machine learning" OR "artificial intelligence" OR "deep learning" OR "predictive modeling") AND ("suicide" OR "self-harm" OR "self- injurious behavior") AND ("adolescent*" OR "teenager*" OR "youth" OR "young people")	1

3.2. Criterios de inclusión y exclusión de estudios

Criterios de inclusión:

- Estudios originales con modelos de ML aplicados a la predicción de conductas autolesivas y suicidas en adolescentes.
- Investigaciones que incluyan muestras de adolescentes (10-19 años).
- Resultados que incluyan métricas de la intervención como sensibilidad, especificidad o precisión predictiva.

Criterios de exclusión:

- Estudios con muestras exclusivas de adultos o poblaciones no adolescentes.
- Publicaciones sin metodología descrita o sin resultados, cartas a director, editoriales o artículos de opinión.
- Texto completo del artículo no disponible.

3.3. Clasificación y filtrado de documentos: declaración PRISMA

El proceso de selección de artículos se ha llevado a cabo en las siguientes fases:

1. Definición de preguntas y criterios.
2. Búsqueda en las fuentes de información.
3. Eliminación de duplicados.
4. Selección de artículos por título y resumen: se evaluó su pertinencia según los criterios de inclusión y exclusión.
5. Revisión del texto completo: Los artículos seleccionados se analizaron en profundidad para confirmar su inclusión.
6. Extracción de datos a tabla de resultados.
7. Evaluación de la calidad de los estudios seleccionados (lectura crítica).
8. Redacción del resumen de resultados.

El proceso de selección de artículos siguió las recomendaciones del PRISMA Statement:

1. Diagrama de flujo PRISMA (Page et al., 2021): Describe gráficamente los pasos para la selección de publicaciones, en las bases de datos utilizadas, desde la búsqueda inicial hasta las consiguientes depuraciones que den como resultado los artículos a incluir en el reporte final.

Tabla 2. Diagrama de flujo PRISMA

DIAGRAMA DE FLUJO PRISMA



3.4. Plan de trabajo

El trabajo de revisión sistemática se ha distribuido en el siguiente cronograma (tabla 3), que refleja un trabajo de 9 semanas, para realizar seis fases:

Tabla 3. Cronograma.

Tarea	Seamanas (inicio 1 de abril - final 2 de junio)								
Definición del protocolo									
Estrategia de búsqueda									
Clasificación y filtrado									
Lectura crítica									
Redacción y discusión de resultados									
Preparación de presentación									

3.5. Sistema de extracción de datos y lectura crítica

Se analizaron las siguientes variables:

- Tipo de estudio: Diseño experimental, observacional o revisión.
- Ámbito: Contexto clínico, educativo o comunitario, sociológico, país, etc.
- Población: Adolescente (10-19 años), tamaño muestral.
- Instrumentos de medida: Algoritmos de ML utilizados, métricas reportadas (sensibilidad, especificidad).
- Variables predictoras.
- Factores de riesgo destacados.
- Limitaciones: Sesgos, calidad metodológica, validez externa.

Se utilizó como apoyo el Checklist PRISMA Statement hasta la sección de resultados para garantizar la transparencia y rigor metodológico (Anexo 1).

4. Resultados

4.1 Redacción de resultados

Seguidamente, se presentan las principales características y datos extraídos de los estudios incluidos en esta revisión sistematizada sobre la prevención del suicidio en adolescentes mediante ML o IA.

La búsqueda sistematizada permitió identificar 24 estudios dentro de los criterios de inclusión predefinidos. Estos trabajos abordaron el uso de tecnología de ML y, en menor medida, de IA, aplicada a la predicción de conductas autolesivas y suicidas en adolescentes.

Los artículos seleccionados se publicaron entre 2021 y 2025, reflejando un interés creciente en esta línea de investigación.

La mayoría de los trabajos se realizaron en Asia, especialmente en Corea del Sur y China, y en menor medida se hallaron publicaciones de Noruega, Alemania, Italia y Estados Unidos.

La tabla 4 resume las características clave de cada estudio (Anexo 2), incluyendo el diseño metodológico, la población analizada, las intervenciones preventivas evaluadas, y los principales hallazgos en términos de efectividad.

Esta síntesis permite una visión general estructurada de la evidencia disponible y facilita la comparación entre las distintas estrategias preventivas.

Tabla 4. Resumen de resultados.

Nº referencia	Referencia (autores, año, título, revista).	País	Diseño del estudio	Objetivo principal	Tamaño muestral	Edad media/rango de edad	Tipo de conducta analizada	Fuente de datos	Tipo de modelo ML/IA usado	Variables predictoras	Eficacia del modelo predictivo	Principales resultados	Factores de riesgo destacados	Limitaciones del estudio	Evaluación de calidad
36	Haghian EF, Liang B, Czajkowski N (2023). Are self-reported suicidal thoughts and suicidal ideation a risk group? Evidence for "true alarm" in a population-representative longitudinal study of Norwegian adolescents. Frontiers in Psychology	Noruega	Estudio longitudinal retrospectivo	Evaluar si los hallazgos positivos en modelos de predicción de suicidio son un grupo de riesgo real	2.505 adolescentes	Edad media 17 años; rango 16–19 años	Intento de suicidio e intento de suicidio	Encuestas/autobioinformes longitudinales	Random forest	Síntomas depresivos, problemas de conducta, estrés familiar, victimización, riesgo real aumentado	AUC 0.80; sensibilidad 0.86; especificidad 0.80	Los hallazgos positivos identificados por el modelo de machine learning en adolescentes con autointerferencia con riesgo real aumentado de intento de suicidio	Depresión, victimización, apoyo social, estrés parental	Subregistro de intentos de suicidio y dependencia de autointerferencias	Calidad moderada; riesgo de sesgo bajo, preocupación moderada en aplicabilidad (autointerferencias)
24	Haghian EF, Bang Nes R, Obadi M, Qin P, Sørensen LL, Bakstun M, Liang B, Czajkowski N (2024). Predictive Value of Self-Reported Suicidal Thoughts and Risk Factors Using Multiple Machine Learning Algorithms. Journal of Youth and Adolescence	Noruega	Estudio longitudinal prospectivo	Analizar factores de riesgo y protección de suicidio adolescente usando machine learning	2.505 adolescentes	Edad media 17 años; rango 16–19 años	Ideación suicida e intento de suicidio	Encuestas/autobioinformes longitudinales	Random forest, XGBoost, redes neuronales profundas (DNN)	Síntomas depresivos, autointerferencia, victimización, estrés parental	AUC 0.81–0.88 (DNN); sensibilidad 0.80; especificidad 0.72–0.78	Modelos de machine learning alcanzaron alta precisión (AUC > 0.80) integrando factores de riesgo y protección	Síntomas depresivos, percepción de bajo apoyo social, estrés parental	Datos incompletos y posibles sesgos de recuerdo	Calidad alta; bajo riesgo de sesgo, buena aplicabilidad en aplicabilidad
38	Haroz EE, Kilbren C, Nestadt PS, Wilcox HC, DeVyster JE, Kharratz H (2021). Using the predictive value of screening to the use of electronic health record data for detecting future suicidal thoughts and behavior in an urban pediatric emergency department: A preliminary analysis. Suicide and Life-Threatening Behavior	Estados Unidos	Estudio retrospectivo	Comparar cribado clínico versus datos de registros electrónicos para predecir riesgo suicida	1.521 pacientes pediátricos	Edad media no reportada; rango 10–18 años (pediátrico)	Ideación suicida y comportamiento suicida	Registros de salud (EHR) y transcripciones de crisis de salud	Random forest	Diagnósticos de salud mental, veías de urgencias, hospitalizaciones previas	AUC 0.84 (screening) vs 0.80 (EHR)	Screening clínico tradicional fue más preciso que EHR para predecir riesgo de suicidio	Diagnóstico psiquiátrico previo, número de visitas hospitalizaciones recientes	Pequeño tamaño muestral y limitaciones en la calidad de los datos del EHR	Calidad moderada; riesgo de sesgo moderado por calidad de datos EHR
41	Cohen J, Wright-Bergman J, Rohle L, Wright D, Campbell M, Grogan D, Smit D, Preston J (2020). A Feasibility Study Using a Machine Learning Suicide Risk Prediction Model Based on Open-Ended Interview Language in Adolescent Therapy Sessions. International Journal of Environmental Research and Public Health	Estados Unidos	Estudio retrospectivo	Evaluar la viabilidad de un modelo de predicción de riesgo suicida basado en lenguaje espontáneo	196 sesiones de terapia	Edad media no reportada; rango 12–17 años	Ideación suicida	Grabaciones de sesiones terapéuticas (transcripciones)	Análisis de lenguaje natural (NLP) y aprendizaje automático supervisado	Características del lenguaje en terapia: ansiedad, autointerferencia negativa	AUC 0.85; precisión 0.83; F1 score 0.81	El lenguaje espontáneo de los adolescentes con ideación suicida con alta precisión (AUC 0.85)	Expresión verbal de desesperanza, lenguaje sobre baja autoestima	Muestra pequeña, falta de validación externa	Calidad baja; alto riesgo de sesgo moderado por calidad de validación externa
25	Czyz EK, Koo HJ, Al-Dajani N, King CA, Nahum-Shani I (2021). Predicting short-term suicidal ideation and self-harm using machine learning: Developing decision tools to identify high-risk adolescents. Psychological Medicine	Estados Unidos	Estudio prospectivo	Predicir pensamientos suicidas y hospitalización en adolescentes	110 adolescentes hospitalizados	Edad media 15.7 años; rango 12–17 años	Ideación suicida	Autointerferencias hospitalización	Random forest y algoritmos de regresión logística penalizada (lasso)	Ánimo depresivo, desesperanza, impulsividad diaria, eventos vitales	AUC 0.70; sensibilidad 0.74; especificidad 0.63	Los cambios diarios en el estado emocional de los adolescentes con ideación suicida a corto plazo	Desesperanza diaria, bajo afecto positivo, alta impulsividad	Limitada generalización fuera de muestra hospitalaria específica	Calidad moderada-alta; bajo riesgo de sesgo moderado por calidad de validación externa
32	Gan Y, Kuang L, Xu X-M, Al-M, He J-H, Wang W, Hong S, Chen J, Cao J, Zhang Q (2025). Research on prediction model of adolescent suicide and self-harm behavior based on machine learning algorithm. Frontiers in Psychiatry	China	Estudio transversal	Desarrollar un modelo de predicción de conducta suicida y autolesiones en adolescentes usando machine learning	19.818 adolescentes	Edad media no reportada; rango 12–18 años	Autolesiones no suicidas e intento de suicidio	Autointerferencias de adolescentes (cuestionarios escolares)	Random forest	Síntomas emocionales, problemas de conducta, apoyo social	AUC 0.73 (autointerferencias); 0.72 (escolares)	El modelo predijo de forma aceptable tanto autolesiones como intentos de suicidio en adolescentes (AUC ~0.73)	Depresión, conflictos familiares, bajo rendimiento académico	Autointerferencias, sesgados y falta de evaluación objetiva	Calidad moderada; riesgo de sesgo moderado por calidad de validación externa
34	Haghian EF, Czajkowski NQ, von Soest T (2023). Predicting suicide attempt among Norwegian adolescents without using suicide-related items: A machine learning approach using longitudinal data. Journal of Affective Disorders	Noruega	Estudio longitudinal prospectivo	Predicir intentos de suicidio sin utilizar ítems directamente relacionados con suicidio	2.505 adolescentes	Edad media 17 años; rango 16–19 años	Intento de suicidio	Encuestas/autobioinformes longitudinales	Random forest	Factores escolares, síntomas emocionales, problemas familiares	AUC 0.80; sensibilidad 0.65	Es posible predecir intentos de suicidio sin usar directamente ítems de suicidio, aunque con algo menor precisión	Problemas emocionales y baja percepción de competencia escolar	Uso limitado de variables directas de autointerferencia	Calidad moderada; riesgo bajo de sesgo, pero aplicabilidad limitada
37	Haghian EF, Czajkowski N, Walby FA, Qin P, Liang B (2024). Suicide attempt risk predicts inconsistent self-reported suicide attempt: A machine learning approach using longitudinal data. Journal of Affective Disorders	Noruega	Estudio longitudinal retrospectivo	Explorar la inconsistencia en los autoinformes de intentos de suicidio en adolescentes	2.505 adolescentes	Edad media 17 años; rango 16–19 años	Intento de suicidio	Encuestas/autobioinformes longitudinales	Random forest	Historia de autolesiones, síntomas emocionales, comportamiento antisocial	AUC 0.83; precisión 0.75	Los adolescentes que negaban intentos pasados pero eran predichos como de alto riesgo, efectivamente tenían factores de riesgo elevados	Historia de autolesiones, depresión, relaciones con padres	Errores de memoria en el autoinforme de intentos de suicidio	Calidad alta; bajo riesgo de sesgo y buena aplicabilidad
45	Xiang Yin, Dan Ma, Kejing Zhu, Dayun Li (2021). Identifying intentional injuries among children and adolescents based on Machine Learning. PLOS ONE	China	Estudio transversal retrospectivo	Detectar lesiones intencionales en adolescentes usando modelos de machine learning	6.750 lesiones registradas (niños y adolescentes)	Edad media no reportada; niños y adolescentes (mayoría 10–18 años)	Autolesiones y violencia (no centrado exclusivamente en suicidio)	Sistema de vigilancia de lesiones (registros hospitalarios)	Naive Bayes, Decision Trees, Random Forest, Adaboost, Deep Neural Networks (DNN)	Características de la lesión, antecedentes de violencia, factores demográficos	AUC 0.89–0.93 según modelo (DNN mejor)	DNN y Adaboost fueron los mejores modelos para identificar lesiones intencionales en adolescentes	Antecedentes de violencia, lesiones recurrentes, violencia escolar	Dificultad para clasificar intenciones de lesiones con precisión total	Calidad moderada-alta; bajo riesgo de sesgo y buena estabilidad entre modelos

José Manuel Granada López

Revisión sistemática sobre el uso de machine learning (ML) para la predicción de conductas autolesivas y suicidas en adolescentes

Nº referencia	Referencia (autores, año, título, revista)	País	Diseño del estudio	Objetivo principal	Tamaño muestral	Edad media/rango de edad	Tipo de conducta analizada	Fuente de datos	Tipo de modelo ML/IA usado	Variables predictoras	Eficacia del modelo predictivo	Principales resultados	Factores de riesgo destacados	Limitaciones del estudio	Evaluación de calidad
33	Jian Yang, Yingjie Chen, Gangyi Yao, Zheng Wang, Xi Bai, Yueqiang Tian, Yongliu Li (2022). Key factors selection on adolescents with non-suicidal self-injury: A support vector machine based approach. Frontiers in Public Health	China	Estudio transversal	Identificar factores asociados a autolesiones no suicidas en adolescentes mediante Support Vector Machine	1.228 adolescentes	Edad media 15,6 años; rango 13–18 años	Autolesiones no suicidas	Cuestionarios psicológicos clínicos	SVM (Support Vector Machine) optimizado	Niveles de ansiedad, depresión, personalidad, traumas en la infancia	Precisión 84%; sensibilidad 82%; especificidad 85%	SVM logró buena predicción de autolesiones a partir de factores emocionales y de personalidad	Ansiedad elevada, rasgos de personalidad borderline, abuso emocional infantil	Muestra limitada a un solo hospital y sesgo cultural	Calidad moderada; riesgo moderado de sesgo (única muestra hospitalaria)
44	Oton Weiler, Luke Sagers, Carl Hanson, James Wang, Quinn Bell, E. Sharron Tass (2021). Predicting suicidal behavior among adolescents using the risk and protective factor framework: A large-scale machine learning approach. PLoS ONE	Estados Unidos	Estudio transversal analítico	Predicir pensamientos y comportamientos suicidas en adolescentes usando factores de riesgo y protección	179.000 adolescentes	Edad media 15,1 años; rango 12–18 años	Ideación suicida y comportamiento suicida	Encuesta "Cross-sectional That Care" más datos censales	Random forest	Consumo de drogas, relaciones familiares, victimización escolar, salud mental	Precisión global 91%; recall 85%	Combinación de factores emocionales y escolares predijo STB con 91% de precisión	Conflictos familiares, apoyo escolar, consumo de drogas	Limitaciones en la actualización de datos censales y subinformes	Calidad alta; bajo riesgo moderado de predicción comunitaria
35	C.L. van Vuren, K. van Mens, D. de Beurs, J. Lokenbol, M.F. van der Wal, P. Cuijpers, M.J.M. Chinapaw (2021). Comparing machine learning to a rule-based approach for predicting suicidal behavior among adolescents: Results from a longitudinal population-based survey. Journal of Affective Disorders	Países Bajos (Holanda)	Estudio longitudinal prospectivo (análisis secundario de cohorte)	Comparar machine learning con reglas simples para predecir comportamiento suicida futuro en adolescentes	1.681 estudiantes	Edad media 14,5 años; rango 13–16 años	Ideación suicida e intento de suicidio	Encuestas longitudinales de programas de salud pública	Random forest y Lasso regression	Historia de ideación suicida, problemas emocionales	AUC 0.79 (Random forest); AUC 0.64 (Lasso); AUC 0.64 (regla clínica)	Machine learning mejoró moderadamente la predicción sobre reglas clínicas tradicionales	Ideación suicida previa, depresión	Pérdida de participantes entre subinformes limitados	Calidad moderada; riesgo moderado de sesgo (subinformes)
5	Raymond Su, James Rufus John, Pimpi Lin (2020). Using machine learning for prediction for self-harm and suicide attempts in adolescents. Psychiatry Research	Australia	Estudio longitudinal prospectivo (análisis secundario de cohorte)	Predicir autolesiones y tentativas de suicidio en adolescentes usando machine learning	2.909 adolescentes	Edad media 15 años; rango 14–17 años	Autolesiones no suicidas e intento de suicidio	Datos longitudinales de subinformes (autobormes)	Random forest	Factores emocionales, depresión, estrés parental, relaciones escolares	AUC 0.74 (autolesiones); AUC 0.72 (suicidio)	Predicción aceptable de autolesiones y suicidio en adolescentes (AUC 0.72-0.74)	Sentimientos de desesperanza, bajo apoyo social, problemas escolares	Dependencia de subinformes y sesgo de memoria	Calidad moderada; riesgo de sesgo bajo, pero buena aplicabilidad moderada
42	Cheng Su, Robert Aseltine, Riddhi Doshi, Kun Guan, Steven C. Foggart, Fai Wang (2020). Predicting suicidal behavior in children and adolescents with electronic health records. Translational Psychiatry	Estados Unidos	Estudio observacional de cohorte	Desarrollar modelos predictivos de riesgo suicida en niños y adolescentes usando registros electrónicos de salud	41.721 adolescentes	Edad media 15 años; rango 10–18 años	Intento de suicidio	Registros electrónicos de salud (EHR)	Random forest, Lasso, XGBoost	Diagnósticos médicos, depresión, ideación, resultados de laboratorio	AUC 0.81–0.86 (autolesiones); AUC 0.81 (intento de suicidio)	EHRs rutinarios permitieron construir modelos de riesgo suicida (AUC > 0.80)	Traumas depresivos, visitas a emergencias, historia psiquiátrica	Falta de variables psicosociales relevantes en el EHR	Calidad alta; bajo riesgo de sesgo, buena aplicabilidad a entornos clínicos
12	Michaela Servi, Silvia Chiaro, Elina Musci, Giovanni Catalini, Alberta Meru, Yany Volpe, Alessandra Biondi, Maria Grazia Di Lorenzo (2020). Machine learning to identify risk factors for suicide in children and adolescents. Science Progress	Italia	Estudio retrospectivo	Identificar predictores tempranos de riesgo suicida en niños y adolescentes hospitalizados	237 adolescentes hospitalizados	Edad media 15,2 años; rango 8–18 años	Ideación suicida, intento de suicidio, autolesiones	Historias clínicas hospitalarias	Redes neuronales artificiales (ANN)	Edad, diagnóstico médico, motivo de ingreso al hospital, intento de suicidio	AUC 0.85	Identificación de factores de riesgo permitió predecir riesgo suicida con AUC de 0.85	Diagnóstico de trastorno de estado de ánimo previo a la ideación suicida	Datos hospitalarios limitados, posible sesgo de selección	Calidad moderada; riesgo moderado de sesgo (datos hospitalarios)
48	Mos Paykin, Oskar Pijpers, Merlijn Ballender, Mos Koenig, Jennifer Nibbelk, Clara Heitner, Olivia Qila, John Bureghiel (2024). Comparison between clinician and machine learning prediction in a randomized controlled trial for nonsuicidal self-injury. BMC Psychiatry	Suecia	Ensayo clínico aleatorizado secundario (análisis retrospectivo)	Comparar la predicción clínica con machine learning en la remisión de autolesiones no suicidas en adolescentes	66 adolescentes	Edad media 15,8 años; rango 13–17 años	Autolesiones no suicidas	Datos de ensayo clínico (evaluaciones clínicas y subinformes)	Modelos predictivos basados en regresión logística y machine learning	Historia de autolesiones, síntomas depresivos, impulsividad	AUC 0.76	La predicción basada en machine learning mejoró ligeramente sobre la predicción clínica en autolesores	Conducta de autolesión previa, desesperanza, impulsividad	Tamaño muestral pequeño y falta de poder estadístico suficiente	Calidad baja; riesgo alto de sesgo (tamaño pequeño)
11	Cheng Zhang, Yu Heu, Daoping Liu, Jun Jin, Shiao Sacco, Kun Chen, Robert Bellini, Fai Wang (2024). Accuracy and transportability of machine learning models for suicidal ideation prediction with longitudinal clinical records. Translational Psychiatry	Estados Unidos	Estudio retrospectivo multicéntrico	Evaluar la precisión y transportabilidad de modelos de predicción de suicidio en adolescentes hospitalizados en clínicas longitudinales	285.320 pacientes	Edad media no reportada; población general adolescente	Intento de suicidio	Registros clínicos longitudinales de múltiples hospitales	Random forest, LightGBM y redes neuronales	Diagnósticos clínicos, visitas hospitalarias, historia psiquiátrica	AUC 0.81–0.86; sensibilidad 53–62%; especificidad 80%	Modelos lograron buena predicción de riesgo a corto y largo plazo con AUC hasta 0.86	Diagnóstico de depresión, antecedentes de hospitalización psiquiátrica	Falta de datos detallados sobre gravedad clínica y tratamientos	Calidad alta; bajo riesgo de sesgo, alta aplicabilidad
39	Lim J.S., Yang C.M., Baek J.W., Lee S.Y., Kim B.N (2022). Prediction Models for Suicide Attempts among Adolescents Using Machine Learning Techniques. Clinical Psychopharmacology and Neuroscience	Corea del Sur	Estudio retrospectivo basado en encuesta	Desarrollar modelos de predicción de intentos de suicidio en adolescentes coreanos	469.482 adolescentes	Edad media 15 años; rango 12–18 años	Intento de suicidio	Encuestas escolares nacionales	Random forest	Síntomas depresivos, eventos vitales estresantes, conductas de riesgo	AUC 0.79	Variables simples del entorno escolar predijeron intentos razonablemente (AUC 0.79)	Depresión, historial de autolesiones, eventos vitales estresantes	Diseño retrospectivo y riesgo de sesgo de subinforme	Calidad moderada; riesgo de sesgo moderado (datos subinformes)

Revisión sistemática sobre el uso de machine learning (ML) para la predicción de conductas autolesivas y suicidas en adolescentes

Nº referencia	Referencia (autores, año, título, revista)	País	Diseño del estudio	Objetivo principal	Tamaño muestral	Edad media/rango de edad	Tipo de conducta analizada	Fuente de datos	Tipo de modelo ML/IA usado	Variables predictoras	Eficacia del modelo predictivo	Principales resultados	Factores de riesgo destacados	Limitaciones del estudio	Evaluación de calidad
10	Lee H, Cho JK, Park J, Lee H, Ford G, Boyer L, et al. (2021). Machine Learning-Based Prediction of Suicidality in Adolescents With Allergic Rhinitis: Derivation and Validation in 2 Independent Nationwide Cohorts. Journal of Medical Internet Research	Corea del Sur	Estudio retrospectivo de cohortes	Predicir riesgo suicida en adolescentes con rinitis alérgica usando machine learning	Cohortes de 7.940 y 7.982 adolescentes	Edad media 14–16 años	Ideación suicida	Encuestas nacionales de salud escolar	Redes profundas (DNN)	Severidad de síntomas alérgicos, calidad de vida, estado emocional	AUC 0.83–0.85	Factores de riesgo alérgicos significativamente asociados con riesgo suicida en adolescentes	Severidad de síntomas de alergia, bajo bienestar emocional	Datos autoinformados y falta de mediciones clínicas objetivas	Calidad alta; bajo riesgo de sesgo, buena aplicabilidad clínica
40	Kim KW, Lim JS, Yang CM, Jang SH, Lee SY (2021). Classification of Adolescent Psychiatric Patients at High Risk of Suicide Using the Personality Assessment Inventory by Machine Learning. Psychiatry Investigation	Corea del Sur	Estudio retrospectivo	Clasificar adolescentes psiquiátricos de alto riesgo suicida usando la Personality Assessment Inventory y machine learning	923 pacientes adolescentes	Edad media 16,2 años; rango 12–18 años	Riesgo de suicidio (ideación e intento)	Inventarios psicológicos (PAL-Adolescent)	Árboles de decisión (CART) y SVM	Puntuaciones de personalidad, impulsividad, problemas afectivos	AUC 0.81 (mejor modelo)	PAL-Adolescent combinado con machine learning permitió buena discriminación de alto riesgo suicida (AUC 0.81)	Altas puntuaciones en rasgos de personalidad borderline y antisocial	Falta de validación externa y sobreajuste potencial	Calidad moderada; riesgo de sesgo bajo, posible sobreajuste
13	Kim H, Son Y, Lee H, King J, Hammoodi A, Choi Y, Kim HJ, Lee H, Ford G, Boyer L, Kwon R, Woo S, Yon DK (2024). Machine Learning-Based Prediction of Suicidal Thinking in Adolescents by Derivation and Validation in 3 Independent Korean Cohorts: Algorithm Development and Validation Study. Journal of Medical Internet Research	Corea del Sur	Estudio retrospectivo de cohortes	Desarrollar y validar modelos de predicción de pensamiento suicida en adolescentes en tres cohortes internacionales	14.382 adolescentes de tres cohortes	Edad media 14–16 años (según cohortes)	Ideación suicida	Encuestas nacionales de salud escolar	Redes profundas (DNN)	Estado emocional, bullying, problemas familiares	AUC 0.82–0.85 (según cohorte)	Modelos basados en DNN mejoraron ideación suicida con alta precisión en cohortes globales (AUC > 0.82)	Bullying escolar, historia de trauma, aislamiento social	Posibles diferencias culturales entre cohortes internacionales	Calidad alta; bajo riesgo de sesgo y buena aplicabilidad internacional
27	Jankovský K, Steger D, Schroeders U (2024). Predicting Lifetime Suicide Attempts in a Community Sample of Adolescents Using Machine Learning Algorithms. Assessment	Alemania	Estudio transversal	Predicir intenciones de suicidio a lo largo de la vida en una muestra comunitaria de adolescentes usando machine learning	1.167 adolescentes	Edad media 15,2 años; rango 13–18 años	Intento de suicidio	Encuestas escolares subadministradas	Random forest, SVM (Support Vector Machine)	Problemas de conducta, depresión, relaciones familiares	AUC 0.77 (Random Forest); AUC 0.75 (SVM)	Random forest y SVM mostraron predicciones moderadas (AUC 0.75–0.77) en población comunitaria	Depresión, estrés escolar, conflictos familiares	Tamaño muestral moderado y falta de replicación externa	Calidad moderada; riesgo de sesgo bajo, pero falta de validación externa
20	Haroz EE, Rebman P, Goldish N, Garcia M, Sule R, Maggio D, Clattenburg E, Mega J, Adams R (2024). Performance of Machine Learning Suicide Risk Models in an American Indian Population. JMIR Network Open	Estados Unidos	Estudio retrospectivo	Evaluar la precisión de modelos de machine learning para riesgo de suicidio en población indígena americana	1.595 (jóvenes navegos)	Edad media 14,8 años; rango 12–18 años	Ideación suicida y comportamiento suicida	Encuestas escolares y registros electrónicos de salud	Random forest y XGBoost	Diagnósticos de salud mental, historia de hospitalización, medicamentos	AUC 0.81–0.83; sensibilidad 0.72; especificidad 0.75	Los modelos de machine learning mostraron alta eficacia predictiva en población indígena (AUC ~0.82)	Historias de suicidio previos, ansiedad clínica, historia de trauma	Falta de generalización fuera de la población indígena estudiada	Calidad moderada; riesgo de sesgo bajo, aplicabilidad limitada fuera de población indígena
43	Mo L, Li H, Zhu T (2022). Exploring the Suicide Mechanism Path of High-Suicide-Risk Adolescents—Based on Weibo Text Analysis. International Journal of Environmental Research and Public Health	China	Estudio transversal con análisis retrospectivo de textos	Explorar mecanismos de suicidio de adolescentes de alto riesgo a partir de textos en redes sociales	1.017 publicaciones de Weibo	Edad media no aplicable; análisis de textos de adolescentes	Ideación suicida	Publicaciones en redes sociales (Weibo)	Word2Vec + random forest	Contenido emocional y desesperanza en redes sociales	AUC 0.82	Análisis de redes sociales mostró alto valor predictivo para ideación suicida (AUC 0.82)	Tono emocional negativo en redes sociales, expresiones de desesperanza	Sesgo de selección en publicaciones de Weibo y falta de verificación clínica	Calidad baja; riesgo alto de sesgo por selección de redes sociales

4.2. Análisis crítico de resultados

Características generales de los estudios incluidos

Los artículos muestran diseños diversos, los hay con metodología transversal, estudios observacionales que analizan datos en un único punto temporal, describiendo prevalencias o relaciones entre variables sin inferencias temporales: (Gan et al., 2025), (Jankowsky et al., 2024), (Yang et al., 2022).

Otros estudios son longitudinales prospectivos, se trata de diseños observacionales que siguen a un grupo de sujetos a lo largo del tiempo, comenzando en el presente y observando eventos o cambios futuros. Permiten establecer secuencias temporales y posibles relaciones causales: (Su et al., 2023), (Haghish et al., 2024a), (Czyz et al., 2023), (Haghish et al., 2023a), (van Vuuren et al., 2021).

Los hay longitudinales retrospectivos, que analizan datos ya existentes recogidos en el pasado para estudiar la evolución de una cohorte a lo largo del tiempo. No tienen control sobre la recogida inicial de datos, pero son útiles para detectar patrones históricos. Se encontraron dos artículos (Haghish et al., 2023b), (Haghish et al., 2024b).

También estudios retrospectivos (no longitudinales o no especificados), observacionales que analizan eventos pasados, sin necesidad de seguimiento temporal. Se basan en registros previos y pueden identificar asociaciones, aunque tienen limitaciones para inferir causalidad. Se identificaron en este epígrafe: (Zang et al., 2024), (Servi et al., 2023), (Haroz et al., 2024), (Zalsman et al., 2016), (Haroz et al., 2021), (Kim et al., 2021), (Cohen et al., 2020).

Otros estudios seleccionados son retrospectivos de cohortes, en los que se identifican cohortes a partir de datos pasados y se evalúan los desenlaces según exposiciones previas. Permiten estimar riesgos relativos, y se hallaron: (Lee et al., 2024), (Kim et al., 2024), (Su et al., 2020).

Tres estudios transversales analíticos o retrospectivos, con análisis comparativos entre grupos o con enfoque retrospectivo (textos o datos archivados), lo que permite hipótesis exploratorias más profundas que un diseño puramente descriptivo. Se incluyeron: (Mo et al., 2022), (Weller et al., 2021), (Yin et al., 2021).

Un artículo exponía un ensayo clínico aleatorizado secundario (análisis retrospectivo), que es el reanálisis retrospectivo de datos obtenidos en un ensayo clínico aleatorizado

previo. Permite responder nuevas preguntas o explorar subgrupos con alto nivel de evidencia, aunque limitado por la hipótesis original del estudio. Se encontró el estudio de Pontén et al. (Pontén et al., 2024).

Se observó una gran heterogeneidad en los tamaños muestrales, que oscilaron entre muestras muy reducidas, como los 66 adolescentes del estudio (Pontén et al., 2024), hasta grandes bases de datos poblacionales, como los 285.320 pacientes del estudio (Zang et al., 2024) o los 468.482 adolescentes del estudio (Lim et al., 2022). Algunos estudios se basan en muestras de alrededor de 2.500 sujetos (Haghish et al., 2024a; Haghish et al., 2023a; Haghish et al., 2023b; Haghish et al., 2024b), mientras que otros alcanzaron cifras más amplias, como los 19.818 adolescentes del estudio (Gan et al., 2025) o los 41.721 del estudio (Su et al., 2020).

Cabe destacar aquellos que no emplearon personas como unidad de análisis, sino otro tipo de datos: el estudio (Cohen et al., 2020) analizó 196 sesiones de terapia como unidad principal, y el estudio (Mo et al., 2022) que basó su análisis en 1.017 publicaciones de Weibo, una red social china. También son reseñables los estudios que utilizaron cohortes múltiples (Lee et al., 2024), que incluyó dos cohortes distintas de 7.940 y 7.982 adolescentes, y (Kim et al., 2024) el que trabajó con 14.362 adolescentes divididos en tres cohortes. En conjunto, esta diversidad metodológica refleja tanto la variedad de enfoques como la amplitud poblacional de las investigaciones revisadas.

La mayoría de los estudios incluidos en la revisión se centraron en población adolescente, con edades medias que oscilaron entre los 14 y los 17 años. Se observaron varios trabajos con edades medias coincidentes en 17 años (Haghish et al., 2024a), (Haghish et al., 2023a), (Haghish et al., 2023b), (Haghish et al., 2024b), todos ellos con rangos muy similares (16-19 años), representando a adolescentes mayores. Otros estudios presentaron edades medias ligeramente menores con 14,5 años (van Vuuren et al., 2021) o con 14,8 años (Haroz et al., 2024), lo que sugiere el interés y la importancia de prevenir el suicidio desde las fases más tempranas de la adolescencia.

Varios trabajos destacaron por su amplitud de rango etario, incluyendo tanto niños como adolescentes: por ejemplo, el estudio (Haroz et al., 2021) incluyó sujetos entre 10 y 18 años, y el (Servi et al., 2023) entre 8 y 18 años. También aparecieron estudios que no reportaron media de edad

(Cohen et al., 2020), centrado en sesiones de terapia con adolescentes, y (Zang et al., 2024) que utilizó una amplia población adolescente sin desgloses por edad. Por último, el estudio (Mo et al., 2022) resultó particular, ya que no tenía media de edad aplicable por basarse en el análisis de publicaciones publicadas en red social por adolescentes, no en sujetos evaluados directamente.

Tipo de conducta relacionada con el suicidio

Los estudios incluidos en la revisión abordaron una variedad de conductas relacionadas con el suicidio y la autolesión en población adolescente. El intento de suicidio fue la conducta más estudiada de forma aislada, como se observó en (Zang et al., 2024), (Jankowsky et al., 2024), (Haghish et al., 2023a), (Haghish et al., 2023b), (Haghish et al., 2024b), (Lim et al., 2022), (Su et al., 2020), mientras que la ideación suicida apareció como variable principal en varios otros (Lee et al., 2024), (Kim et al., 2024), (Haroz et al., 2024), (Cohen et al., 2020), (Mo et al., 2022). Algunos trabajos analizaron conjuntamente ideación e intento de suicidio (Haghish et al., 2024a), (van Vuuren et al., 2021), (Kim et al., 2021), proporcionando así una perspectiva más completa del fenómeno del suicidio en conjunto.

También destacaron aquellos que incluyeron autolesiones no suicidas, bien como foco principal (Yang et al., 2022), (Pontén et al., 2024), o combinadas con intento de suicidio (Su et al., 2023), (Gan et al., 2025). Un caso particular fue el estudio (Servi et al., 2023), que abordó de forma simultánea la ideación, el intento y la autolesión, proporcionando una visión integral de la conducta autolesiva. Por otro lado, el estudio (Yin et al., 2021) incluyó tanto autolesiones como violencia, pero no se centró exclusivamente en suicidio, lo que lo diferencia metodológicamente del resto.

En suma, aunque predominó el análisis del intento de suicidio, muchos estudios optaron por abordar múltiples dimensiones del espectro suicida y autolesivo, lo que evidencia un enfoque más amplio y contextualizado del fenómeno.

Modelos de ML utilizados

Los estudios incluidos en la revisión emplearon una amplia variedad de modelos de ML para predecir conductas suicidas o autolesivas, siendo el algoritmo random forest el más utilizado, ya sea de forma exclusiva o combinado con otros métodos. Destacaron estudios que usaron random forest de manera aislada sobre datos de encuestas o autoinformes longitudinales (Haghish et al., 2024a), (Haghish et al.,

2023a), (Haghish et al., 2023b), (Haghish et al., 2024b), así como aquellos que lo combinaron con técnicas más avanzadas como XGBoost y redes neuronales profundas (Deep Neural Networks, DNN), como en (Gan et al., 2025), (Lim et al., 2022), (Weller et al., 2021).

Algunos trabajos aplicaron modelos más sofisticados o combinaciones complejas. Por ejemplo, el estudio (Yin et al., 2021) aplicó una batería de algoritmos, incluyendo Naive Bayes, Decision Trees, Adaboost, Random Forest y DNN, sobre registros hospitalarios de lesiones. El estudio (Mo et al., 2022), centrado en publicaciones en redes sociales (Weibo), aplicó procesamiento de lenguaje natural mediante Word2Vec combinado con random forest, lo que lo diferenció metodológicamente.

A destacar también, los que aplicaron redes neuronales artificiales (ANN) sobre historias clínicas hospitalarias (Kim et al., 2024), o los que emplearon modelos basados en regresión penalizada (Lasso) junto con random forest, como los estudios (Servi et al., 2023; Su et al., 2020). También se observaron estudios con uso de Support Vector Machines (SVM), tanto de forma aislada (Yang et al., 2022) como combinada (van Vuuren et al., 2021).

En conjunto, la tendencia mostró una preferencia por algoritmos de clasificación de suficiente confianza como random forest, complementados en varios casos con técnicas avanzadas como DNN, SVM o LightGBM, adaptadas al tipo y complejidad del dato disponible.

Fuente de datos utilizada

Los estudios revisados emplearon una amplia gama de fuentes de datos, predominando aquellos que recurrieron a encuestas o autoinformes, en especial de carácter longitudinal (Haghish et al., 2024a), (Haghish et al., 2023a), (Haghish et al., 2023b), (Haghish et al., 2024b), que utilizaron instrumentos autoadministrados repetidos en el tiempo para recoger información subjetiva sobre conducta suicida. También se emplearon cuestionarios escolares o programas de salud pública (van Vuuren et al., 2021), (Lim et al., 2022), (Weller et al., 2021), así como encuestas nacionales de salud escolar (Lee et al., 2024), (Kim et al., 2024), lo que permitió trabajar con grandes bases de datos poblacionales.

Por otro lado, un grupo de estudios se basó en fuentes clínicas u hospitalarias. Algunos emplearon registros electrónicos de salud (EHR), (Haroz et al., 2021), (Su et al., 2020), mientras que otros analizaron historias clínicas hospitalarias

(Servi et al., 2023), o datos longitudinales de hospitales múltiples (Zang et al., 2024). También se incluyeron estudios basados en datos de ensayo clínico (Pontén et al., 2024), lo cual aporta un mayor control sobre las variables clínicas.

Destacaron además algunas fuentes no convencionales por su originalidad. El estudio (Cohen et al., 2020) utilizó transcripciones de sesiones terapéuticas, lo que permitió aplicar análisis de lenguaje natural, y el estudio (Mo et al., 2022) trabajó con publicaciones en redes sociales (Weibo), abriendo nuevas posibilidades de detección de riesgo mediante el análisis del discurso en entornos digitales.

Finalmente, se utilizaron fuentes complementarias como sistemas de vigilancia de lesiones (Yin et al., 2021), inventarios psicológicos estandarizados (Kim et al., 2021) y cuestionarios clínicos validados (Yang et al., 2022), mostrando la diversidad metodológica de la literatura en cuanto a origen y tipo de datos empleados.

Variables predictoras más utilizadas

Los estudios incluidos emplearon una gran diversidad de variables predictoras, principalmente psicológicas, sociales y clínicas, para modelar conductas suicidas o autolesivas en adolescentes. Las variables más comunes fueron los síntomas depresivos, (presentes en artículos 36, 24, 25 y 39), muchas veces acompañados de desesperanza, impulsividad o problemas familiares. También se usaron combinaciones complejas de factores emocionales y conductuales (Gan et al., 2025), que incluyó apoyo social, problemas de conducta y síntomas emocionales.

Algunos estudios se centraron en historia clínica o diagnósticos médicos como predictores (Haroz et al., 2021), incorporando diagnósticos de salud mental y hospitalizaciones, o el (Su et al., 2020), que incluyó medicación y resultados de analíticas. El estudio (Servi et al., 2023) fue más específico, utilizando el diagnóstico psiquiátrico y el método del intento de suicidio como predictores. También fueron frecuentes los antecedentes de autolesión, como en los estudios (Haghish et al., 2024b), (Pontén et al., 2024), lo que da cuenta de su importancia como marcador de riesgo.

Destacaron estudios con variables no tradicionales o innovadoras. Así, el estudio (Cohen et al., 2020) analizó características del lenguaje en sesiones terapéuticas como indicadores de autoimagen negativa y desesperanza. De forma similar, el estudio (Mo et al., 2022) exploró el contenido emocional y la desesperanza en publicaciones en redes

sociales, abriendo nuevas vías para la detección precoz mediante lenguaje digital.

Otros estudios adoptaron un enfoque multidimensional, como el (Weller et al., 2021), que incluyó consumo de drogas, victimización escolar, relaciones familiares y salud mental, o el (Yang et al., 2022), que evaluó ansiedad, depresión, personalidad y traumas infantiles. En conjunto, los trabajos reflejaron la complejidad del fenómeno suicida y la necesidad de abordar múltiples dimensiones emocionales, sociales y clínicas para mejorar la predicción del riesgo.

Eficacia del modelo predictivo

Los estudios revisados emplearon principalmente el área bajo la curva ROC (AUC) como métrica de desempeño para evaluar la eficacia de los modelos predictivos. En general, los valores AUC oscilaron entre 0.70 y 0.93, lo que indica una eficacia moderada a alta. Destacaron varios estudios con desempeño excelente: el estudio (Yin et al., 2021) alcanzó un AUC entre 0.89 y 0.93, siendo el modelo basado en DNN el que obtuvo los mejores resultados. De forma similar, el estudio (Weller et al., 2021) reportó una precisión global del 91% y un recall (detección de positivos) del 85%, mientras que el (Yang et al., 2022) alcanzó una precisión del 84% y especificidad del 85%, reflejando un muy buen rendimiento, detectando correctamente los negativos.

Varios estudios presentaron un desempeño sólido con AUC en torno a 0.80-0.85, como los estudios (Haghish et al., 2023b), (Haroz et al., 2021), (Cohen et al., 2020), (Servi et al., 2023), (Lee et al., 2024), (Kim et al., 2024), (Haroz et al., 2024), (Mo et al., 2022). En particular, el estudio (Cohen et al., 2020), que utilizó características del lenguaje en terapia, obtuvo un AUC de 0.85, precisión del 83% y F1 score de 0.81, lo que demuestra que incluso variables no convencionales pueden ofrecer alto valor predictivo.

Estudios como el (Czyz et al., 2023) o el (Gan et al., 2025) mostraron rendimientos más modestos, con AUC en torno a 0.70-0.73, lo que puede deberse a la complejidad de las variables, la calidad del dato o el diseño del modelo. También es notable el estudio (van Vuuren et al., 2021), que comparó tres enfoques y encontró diferencias relevantes: AUC 0.79 para Random Forest, 0.76 para Lasso y apenas 0.64 para una regla clínica, lo que evidencia la superioridad de los enfoques de aprendizaje automático frente a métodos tradicionales.

En resumen, la mayoría de los modelos alcanzaron una efectividad adecuada, y varios lo hicieron de forma sobre-

saliente. El uso combinado de múltiples métricas (AUC, precisión, sensibilidad, especificidad) permitió una valoración más matizada del rendimiento.

Principales hallazgos

Los estudios revisados mostraron hallazgos relevantes sobre la capacidad predictiva de los modelos de ML aplicados a conductas suicidas y autolesivas. Varios trabajos reportaron una alta precisión en la predicción, destacando el uso de múltiples fuentes de datos y técnicas avanzadas. El estudio (Haghighi et al., 2024a) integró factores de riesgo y protección, logrando AUC superiores a 0.80, mientras que el (Weller et al., 2021) alcanzó una precisión del 91% combinando factores sociales y escolares. Modelos basados en redes neuronales profundas (DNN) también destacaron: el estudio (Yin et al., 2021) mostró que DNN y Adaboost fueron los modelos más eficaces, y el (Kim et al., 2024) obtuvo alta precisión para ideación suicida en cohortes globales.

Algunos trabajos ofrecieron enfoques novedosos. El estudio (Cohen et al., 2020) demostró que el análisis del lenguaje espontáneo en terapia podía predecir riesgo suicida con un AUC de 0.85, y el (Mo et al., 2022) utilizó publicaciones en redes sociales para detectar ideación suicida con alta eficacia (AUC 0.82). De igual modo, el estudio (Lee et al., 2024) identificó como predictores significativos factores médicos aparentemente alejados del campo, como la rinitis alérgica.

También destacaron estudios que, sin alcanzar máximos niveles de precisión, ofrecieron aportaciones valiosas. Por ejemplo, el (Haghighi et al., 2023a) mostró que era posible predecir intentos de suicidio sin emplear ítems directos sobre suicidio, y el (Haghighi et al., 2024b) evidenció que algunos adolescentes negaban intentos previos, pero presentaban alto riesgo según el modelo, validando la utilidad del ML para captar casos ocultos.

Factores de riesgo destacados

Los estudios revisados identificaron una amplia gama de factores de riesgo asociados con conductas suicidas y autolesivas en adolescentes, siendo los más frecuentes la depresión, la desesperanza, el bajo apoyo social o familiar, y los antecedentes de autolesión o trauma. El estudio (Haghighi et al., 2023b) encontró que la victimización escolar, la depresión y el escaso apoyo parental eran predictores relevantes, mientras que el (Haghighi et al., 2024a) destacó el papel del estrés parental y la percepción de bajo apoyo social como factores claves.

Diversos estudios clínicos y basados en registros identificaron predictores más objetivos. Por ejemplo, el estudio (Haroz et al., 2021) halló que diagnósticos psiquiátricos previos, visitas a urgencias y hospitalizaciones recientes estaban vinculados a mayor riesgo, y el (Servi et al., 2023) encontró que los métodos letales previos y los trastornos del estado de ánimo eran factores determinantes. Los estudios (Su et al., 2020), (Zang et al., 2024) confirmaron que los trastornos depresivos y el historial de salud mental eran relevantes desde una perspectiva médica.

Algunos identificaron patrones más sutiles, el estudio (Cohen et al., 2020) reveló que la expresión verbal de desesperanza y baja autoestima en sesiones terapéuticas podía predecir riesgo, y el (Mo et al., 2022) demostró que un tono emocional negativo en redes sociales también era indicador de ideación suicida.

Los factores sociales y escolares también resultaron centrales en varios estudios. El (Weller et al., 2021) relacionó el acoso escolar y los conflictos familiares con el riesgo, y el (Kim et al., 2024) subrayó el aislamiento social, bullying y trauma como elementos clave. Finalmente, variables de personalidad como rasgos borderline o antisociales fueron identificadas como relevantes en los estudios (Yang et al., 2022; Kim et al., 2021), mientras que el (Lee et al., 2024) encontró que incluso síntomas alérgicos severos podían asociarse a malestar emocional y mayor riesgo suicida.

Limitaciones de los estudios

Los estudios incluidos en la revisión presentaron diversas limitaciones metodológicas que afectan la validez y generalización de sus hallazgos. Una de las más comunes fue la dependencia de autoinformes (información aportada subjetivamente por los participantes), lo cual puede introducir sesgos de memoria y subregistro de eventos sensibles como el intento de suicidio (Haghighi et al., 2023b), (Haghighi et al., 2024a), (Su et al., 2023), (Haghighi et al., 2023a). Este problema fue especialmente relevante en el estudio (Haghighi et al., 2024b), que advirtió errores de memoria al reportar intentos pasados, y en el (Lee et al., 2024), donde no se incorporaron mediciones clínicas objetivas.

Varias investigaciones se vieron afectadas por el tamaño muestral reducido (Haroz et al., 2021), (Cohen et al., 2020), (Pontén et al., 2024), lo que limitó el poder estadístico y la posibilidad de generalizar los resultados. El estudio (Czyz et al., 2023), por ejemplo, centrado en pacientes hospitalizados, tuvo proble-

mas de extrapolación a otras poblaciones. También fue común la falta de validación externa o replicación (Kim et al., 2021), (Jankowsky et al., 2024), (Cohen et al., 2020), lo que restringe la robustez de los modelos propuestos.

Otras limitaciones estuvieron relacionadas con la calidad y alcance de las bases de datos clínicas. El estudio (Su et al., 2020) destacó la ausencia de variables psicosociales relevantes en el EHR, mientras que el (Zang et al., 2024) señaló la falta de datos detallados sobre gravedad clínica o tratamientos. Asimismo, el (Servi et al., 2023) evidenció un posible sesgo de selección en los registros hospitalarios utilizados.

Por último, varios estudios enfrentaron limitaciones en la representación poblacional. El estudio (Kim et al., 2024) mencionó posibles diferencias culturales entre cohortes internacionales, el (Haroz et al., 2024) advirtió sobre la falta de generalización fuera de la población indígena, y el (Mo et al., 2022) indicó un sesgo de selección en las publicaciones de Weibo, sin verificación clínica directa.

En conjunto, estas limitaciones invitan a una interpretación cautelosa de los hallazgos y subrayan la necesidad de replicar los estudios en muestras más amplias, clínicas y diversas, incorporando fuentes de datos más completas y validadas.

Evaluación de calidad

Con la finalidad de valorar sistemáticamente el riesgo de sesgo y la aplicabilidad de los estudios incluidos en esta revisión, se utilizó de forma retrospectiva el PROBAST (Prediction model Risk Of Bias Assessment Tool), herramienta destinada a evaluar estudios y validación de modelos predictivos (Anexo 3).

A partir de los datos extraídos de los 24 estudios incluidos en la revisión, se aplicó retrospectivamente la herramienta PROBAST. Destacan algunos datos sobre la calidad de los estudios:

- El 58% de los estudios presentan un riesgo global de sesgo alto, principalmente por el uso exclusivo de autoinformes, muestras pequeñas o falta de validación externa.
- 37% de los estudios presentan un riesgo moderado, generalmente por limitaciones en los predictores o ausencia de seguimiento longitudinal.
- Solo un estudio (4%) fue clasificado con riesgo bajo de sesgo.

En cuanto a la aplicabilidad:

- 10 estudios (42%) presentan aplicabilidad alta, al contar con muestras amplias, buena definición de variables y métodos clínicamente transferibles.
- 12 estudios (50%) presentan una aplicabilidad adecuada, aunque con limitaciones contextuales o culturales.
- Solo 2 estudios (8%) tienen una aplicabilidad limitada, debido a entornos muy específicos o fuentes de datos no replicables.

Estos resultados refuerzan la necesidad de mejorar el diseño y metodología de los estudios y la transparencia en el desarrollo de modelos predictivos basados en ML aplicados a la salud mental, especialmente para su futura integración en ámbitos sanitarios o escolares.

5. Discusión

5.1 Aportaciones a la evidencia disponible

Esta revisión sistematizada representa una aportación de interés al conocimiento acumulado sobre el uso de modelos de aprendizaje automático (ML) para la predicción de conductas autolesivas y suicidas en población adolescente. La síntesis de 24 estudios recientes permite confirmar que estos modelos, lejos de suponer solo una novedad tecnológica, comienzan a consolidarse como herramientas útiles, con una capacidad predictiva que supera a los métodos tradicionales utilizados con anterioridad en salud mental.

Los resultados obtenidos confirman que los modelos ML tienen una capacidad predictiva relevante para identificar adolescentes en riesgo de conducta suicida o autolesiva, corroborando hallazgos previos en la literatura. Por ejemplo, el metaanálisis de Gaynor et al. en 2023 ya apuntaba a la necesidad de diseñar intervenciones más precisas y 31 personalizadas basadas en factores individuales de riesgo, algo que los modelos de ML permiten con mayor eficacia que los enfoques tradicionales (Gaynor et al., 2023).

De igual modo, estudios como los de Su et al. (Su et al., 2023; Su et al., 2020) han demostrado que los algoritmos basados en registros clínicos electrónicos ofrecen una alta sensibilidad para detectar ideación suicida incluso en fases tempranas del malestar psicológico, lo que da idea de su aplicabilidad como herramienta predictora.

De hecho, la presente revisión también confirma que los modelos que integran datos clínicos estructurados y no estructurados (como texto libre o registros longitudinales) alcanzan

valores AUC superiores a 0.85, como en los trabajos de Zang et al. (Zang et al., 2024) y Cansel et al. (Cansel et al., 2023).

Por otro lado, estudios como el de Kim et al. (Kim et al., 2024), que incluyeron múltiples cohortes internacionales, avalan la transferibilidad de estos modelos a distintos contextos geográficos, algo que también se ha observado en esta revisión con trabajos realizados en Asia, Europa y América. No obstante, aún existe una sobre-representación de estudios en entornos sanitarios asiáticos, lo que plantea interrogantes sobre la aplicabilidad global de los resultados a otros países con realidades sociales, culturales y administrativas bien diferentes.

Recientes publicaciones también han advertido sobre las limitaciones éticas y metodológicas del ML en salud mental, especialmente en adolescentes. Smith et al. (2023) señalan el riesgo de colocar etiquetas estigmatizantes y la necesidad de validar estos modelos en contextos reales con el consentimiento informado adecuado y protocolos de intervención bien definidos. La presente revisión también reconoce estas limitaciones e insiste en una aplicación de esta tecnología responsable y regulada (Smith et al., 2023).

Estudios innovadores como el de Mo et al. (2022), que utilizan análisis de lenguaje en redes sociales, y De Luca et al. (2024), con herramientas específicas como el SCI-2, muestran vías alternativas que refuerzan el potencial del ML más allá del contexto clínico, algo que también se refleja en los resultados de esta revisión.

Finalmente, esta revisión sistematiza información dispersa y heterogénea, permitiendo visualizar el estado actual del campo, los algoritmos más utilizados (por ejemplo, random forest, DNN, SVM), las métricas de rendimiento más frecuentes (AUC, sensibilidad, precisión) y las limitaciones recurrentes de los estudios. Este mapa puede servir de guía tanto para investigadores como para responsables de políticas públicas interesados en aplicar estas herramientas a contextos clínicos, escolares o comunitarios.

5.2 Interpretación de los resultados

Los resultados sugieren que los modelos de ML aplicados a la predicción del riesgo suicida en adolescentes presentan un rendimiento satisfactorio, con valores de AUC, sensibilidad y precisión generalmente superiores a 0.80, y no solo son capaces de esta predicción, sino que pueden hacerlo utilizando variables no convencionales, como análisis de texto, patrones de lenguaje o datos de redes sociales. Cuando se

combinan datos clínicos con autoinformes, redes sociales o lenguaje natural, los modelos no solo mejoran su capacidad predictiva, sino que logran captar perfiles de riesgo más sutiles, complejos y cambiantes. Este resultado refuerza la necesidad de diseñar sistemas híbridos de recogida de datos que combinen registros formales e indicadores ecológicos momentáneos, respetando siempre los principios éticos y legales del uso de datos sensibles en menores.

También resulta destacable la utilidad de los algoritmos en la detección de señales no verbales o no explícitas. Varios estudios incluidos en esta revisión demuestran que el ML puede identificar adolescentes en alto riesgo incluso cuando ellos mismos niegan tentativas previas o no reconocen síntomas depresivos. Este hallazgo es importante porque permite captar señales precoces y no explícitas, lo que resulta especialmente útil en adolescentes que suelen ocultar su malestar o minimizar su sufrimiento en contextos clínicos tradicionales, a los que les cuesta acceder, buscando la información y ayuda entre sus iguales.

Aún con todo lo expuesto, los resultados también deben interpretarse con prudencia. El rendimiento del modelo dependerá de la calidad, cantidad y naturaleza de los datos utilizados. Las bases de datos pequeñas, homogéneas o mal estructuradas limitan la capacidad de generalización del modelo. Igualmente, muchos estudios carecen de validación externa o de replicación independiente, lo que obliga a mantener una actitud crítica sobre el entusiasmo excesivo respecto a su aplicabilidad clínica inmediata.

Además, los modelos más eficaces tienden a ser aquellos que integran múltiples capas de información (clínica, conductual, emocional y social), reforzando la necesidad de un enfoque multidimensional en la predicción del riesgo suicida. Esta visión es clave para evitar falsos positivos o el etiquetado estigmatizante ya comentado, complementándose con su capacidad de informar y guiar la intervención, siendo éste uno de los aspectos más prometedores del ML en salud mental infantojuvenil, siempre que se emplee como complemento a la evaluación clínica y no como sustituto.

5.3 Implicaciones clínicas y de investigación

La incorporación de modelos de ML en los sistemas de salud mental permitiría pasar de una prevención reactiva a una prevención proactiva, basada en la identificación temprana de adolescentes en riesgo y en la activación de recursos antes de que ocurran conductas suicidas o autolesivas. Esto es

particularmente relevante en contextos donde los servicios están saturados, el acceso a profesionales de salud mental es limitado, o el estigma impide la demanda explícita de ayuda.

En la práctica, los modelos de ML podrían integrarse en sistemas de historia clínica electrónica para generar alertas automatizadas, o bien formar parte de plataformas digitales diseñadas para el seguimiento emocional de adolescentes. También podrían utilizarse en el ámbito educativo para facilitar un trabajo colaborativo con los equipos de orientación, detectando señales de alerta invisibles a los profesionales docentes. Sin embargo, su implementación exige garantías claras: protocolos de actuación, supervisión profesional, consentimiento informado, derecho a explicación del resultado y garantías de no estigmatización.

Desde el punto de vista de la investigación, esta revisión visualiza la necesidad de estudios con mayor calidad metodológica: diseños longitudinales, muestras representativas y validación externa multicéntrica. También es urgente incorporar enfoques más inclusivos y participativos, en los que adolescentes, familias, educadores y profesionales sanitarios participen en el diseño de los modelos y en la definición de sus criterios de aplicabilidad y utilidad.

Otro reto importante es avanzar en la construcción de modelos explicables y éticamente transparentes. La opacidad de muchos algoritmos de aprendizaje profundo limita su utilidad en contextos sensibles como la salud mental, donde el contexto y causa del resultado es tan importante como el resultado mismo. La investigación debe orientarse a desarrollar herramientas que sean a la vez precisas, comprensibles y auditables, integrando principios de justicia algorítmica y equidad digital.

No obstante, es imprescindible avanzar hacia marcos normativos que regulen el uso ético de estos algoritmos, garanticen la protección de datos sensibles y eviten falsos positivos que puedan dañar la imagen o bienestar del adolescente. Las implicaciones clínicas y científicas van más allá de la predicción: sitúan ante la oportunidad de transformar la forma en que se entiende, evalúa y previene el sufrimiento emocional en la adolescencia.

5.4 Limitaciones de la revisión

A pesar del rigor metodológico aplicado en la revisión sistematizada, se deben señalar algunas de sus limitaciones. En primer lugar, no se realizó metaanálisis debido a la elevada heterogeneidad de diseños, métricas y poblaciones.

Esta variabilidad metodológica dificulta la comparación directa entre estudios y la síntesis estadística agregada.

Como segunda limitación, la revisión puede estar sujeta a sesgo de publicación, ya que los estudios con resultados positivos o con métricas de precisión más elevadas tienden a publicarse con mayor frecuencia que aquellos con resultados negativos o poco concluyentes. Esto puede haber sobredimensionado la aparente efectividad de los modelos de ML.

En tercer lugar, la búsqueda bibliográfica se limitó a artículos publicados en inglés y español, lo que puede haber excluido algún estudio relevante en otros idiomas. Asimismo, se priorizó el acceso a texto completo y publicaciones recientes (últimos 5 años), lo cual, aunque justificado por la rápida evolución del campo, puede haber omitido trabajos fundacionales o aún vigentes.

Otra limitación radica en la falta de validación independiente de los estudios incluidos. Muchos trabajos no fueron replicados en otros contextos, ni emplearon muestras externas, lo que reduce la robustez de los modelos propuestos y su aplicabilidad clínica general. Esta carencia de validación externa se ha identificado como uno de los principales desafíos en el desarrollo de herramientas predictivas basadas en inteligencia artificial en salud mental.

Además, la revisión fue realizada por un solo investigador, lo que incrementa el riesgo de sesgo en la selección e interpretación de los estudios. A pesar del uso de criterios sistemáticos (PRISMA, PROBAB), la revisión no contó con metodologías como doble ciego o verificación cruzada, que aumentarían la fiabilidad en las revisiones sistemáticas.

Por último, no se aplicaron análisis de sensibilidad ni se evaluó en profundidad la calidad de los algoritmos desde una perspectiva técnica (por ejemplo, sobreajuste, validación cruzada, interpretabilidad), lo cual hubiera permitido una evaluación más detallada acerca de la madurez de los modelos revisados.

6. Aplicabilidad y nuevas líneas de investigación

6.1 Aplicabilidad de la tecnología ML

La incorporación de algoritmos de tecnología ML en la predicción del riesgo suicida y conducta autolesiva en adolescentes representa una innovación relevante en el ámbito de la salud mental infanto-juvenil. Su principal aplicabilidad radica en la capacidad de identificar patrones complejos de riesgo a partir de múltiples fuentes de datos: registros clíni-

cos electrónicos, cuestionarios en ámbito escolar, autoinformes psicológicos, e incluso interacciones en redes sociales, facilitando así una detección más temprana y personalizada. En entornos asistenciales, se podrían implementar tecnologías de ML e IA orientadas a dar soporte a intervenciones de cribado, especialmente en atención primaria, salud escolar o unidades de salud mental. Servirían para alertar de riesgo alto de suicidio o de conductas autolesivas en algunas personas, con notificaciones a los profesionales sanitarios de referencia para actuaciones tempranas. También puede resultar útil en el ámbito educativo, su uso en plataformas digitales o aplicaciones de seguimiento de malestar emocional podría ayudar a docentes, orientadores y familias a reconocer señales de alarma de forma preventiva.

Pero esta tecnología se enfrenta a retos éticos y legales, en primer lugar, por no haber una demanda explícita del sujeto o de sus responsables legales para la intervención sanitaria, y por otro lado, se debe garantizar la privacidad y el consentimiento informado en el uso de datos sensibles, minimizar los falsos positivos que puedan estigmatizar a los adolescentes, y asegurar la supervisión profesional de los resultados generados por el modelo.

La aplicabilidad real dependerá de la validación de estos aplicativos en condiciones de práctica habitual, con datos de vida real y observando la diversidad cultural, social y tecnológica.

6.2 Nuevas líneas de investigación propuestas

Con lo aportado por esta revisión sistematizada, se proponen posibles vías para la investigación en un futuro. En primer lugar, se requiere el desarrollo de estudios longitudinales con diseño prospectivo que permitan validar los modelos existentes en muestras representativas y en diferentes contextos socioculturales. Esta validación externa es clave para garantizar la generalización de los resultados.

En segundo lugar, es necesario explorar la integración de variables protectoras: el apoyo familiar, la resiliencia, la participación comunitaria, etc., dentro de los modelos de ML, con la finalidad, además de detectar el riesgo, de identificar factores protectores que puedan ser activados y reforzados desde la intervención.

Como tercera propuesta, se sugiere investigar el uso de datos no estructurados (texto, mensajes de voz, lenguaje natural) mediante técnicas de procesamiento del lenguaje natural (NLP) y análisis emocional, lo cual ha mostrado resultados prometedores, pero aún incipientes.

Por último, se propone investigar el impacto de la implementación real de estos sistemas en entornos clínicos o educativos: cómo afectan a la toma de decisiones, a la relación terapéutica profesional - paciente y al impacto real en la salud mental. Promover la investigación traslacional facilitaría que estudios teóricos o en condiciones casi "de laboratorio" sean implementadas en la práctica habitual.

6.3 Implicaciones sociopolíticas

La integración de tecnologías de IA y el ML en la prevención de conductas autolesivas y suicidas en adolescentes plantea desafíos que trascienden el ámbito sanitario y requieren un abordaje político y social coordinado.

Es fundamental que las políticas de salud mental incluyan el desarrollo y financiación de herramientas tecnológicas predictivas como parte de los programas nacionales de prevención del suicidio, facilitando su uso ético, seguro y equitativo.

En el entorno educativo, la administración debe fomentar la colaboración entre los centros escolares y los dispositivos sanitarios de salud mental infantojuvenil, favoreciendo la formación del profesorado, directivos y orientadores en la identificación de señales de alerta y en el uso responsable de tecnologías de cribado digital. Este esfuerzo debería ir acompañado de campañas de sensibilización que reduzcan el estigma asociado a la enfermedad mental y promuevan una cultura de cuidado, de inclusión, de apoyo y de prevención del malestar emocional.

Asimismo, es imprescindible una normativa clara y actualizada sobre el uso de datos personales en menores, estableciendo mecanismos de consentimiento adaptados a la edad, derechos de acceso y rectificación, y protocolos para evitar usos discriminatorios o punitivos de los resultados generados por algoritmos predictivos.

Finalmente, las políticas públicas deben asegurar la equidad en el acceso a estas herramientas, evitando que las brechas digitales u otras condiciones socioeconómicas refuerzan desigualdades ya existentes en salud mental. La tecnología, bien regulada e implementada con perspectiva de derechos, puede ser un recurso potente para proteger la vida y el bienestar de los adolescentes.

7. Conclusiones

- Los modelos que usan ML representan una herramienta emergente y eficaz para la predicción de conductas autolesivas y suicidas en adolescentes. La capacidad de estos algoritmos para integrar múltiples tipos de datos: registros clínicos estruc-

turados, autoinformes, lenguaje natural o publicaciones en redes sociales, permite mejorar la sensibilidad y especificidad en la identificación de casos de riesgo, superando en muchos estudios los métodos de cribado tradicionales.

- La traslación de estas investigaciones a la práctica clínica y a la prevención todavía tiene dificultades. Las limitaciones metodológicas de los estudios disponibles, como la falta de validación externa, los sesgos derivados del uso de autoinformes o la baja representatividad poblacional, reducen la generalizabilidad de los resultados.
- Se requiere normativa y legislación actualizada y potente que garantice la protección de los datos personales y eviten el estigma asociado a la identificación automatizada del riesgo suicida en menores.

• Los resultados refuerzan la necesidad de seguir desarrollando y validando modelos de ML con diseño multicéntrico, enfoque intercultural y adaptabilidad a la práctica clínica. La integración de estas tecnologías en contextos educativos y de salud mental infantojuvenil, acompañada de estrategias de intervención temprana, puede convertirse en una herramienta clave para reducir una de las principales causas de mortalidad entre adolescentes.

• La colaboración entre profesionales de la salud mental, docentes, desarrolladores y analistas de tecnología y ciencias de datos, juristas y responsables políticos será fundamental para desarrollos futuros de esta tecnología orientada a la prevención del suicidio y autolesiones en adolescentes.

BIBLIOGRAFÍA

- Bazrafshan, M., & Sayehmiri, K. (2024). Predicting suicidal behavior outcomes: An analysis of key factors and machine learning models. *BMC Psychiatry*, 24(World Health Organization, 2014), 841. <https://doi.org/10.1186/s12888-024-06273-2>
- Campillo Serrano, C., & Fajardo Dolci, G. (2021). Prevención del suicidio y conducta suicida. *Gaceta Médica de México*, 157(Su et al., 2023), 547-552. <https://doi.org/10.24875/GMM.M21000611>
- Cansel, N., Yagin, F. H., Akan, M., & Ilkay Aygul, B. (2023). Interpretable estimation of suicide risk and severity from complete blood count parameters with explainable artificial intelligence methods. *Psychiatria Danubina*, 35(World Health Organization, 2014), 62-72. <https://doi.org/10.24869/psyd.2023.62>
- Cohen, J., Wright-Berryman, J., Rohlf, L., Wright, D., Campbell, M., Gingrich, D., et al. (2020). A feasibility study using a machine learning suicide risk prediction model based on open-ended interview language in adolescent therapy sessions. *International Journal of Environmental Research and Public Health*, 17(Bazrafshan y Sayehmiri, 2024), 8187. <https://doi.org/10.3390/ijerph17218187>
- Czyz, E. K., Koo, H. J., Al-Dajani, N., King, C. A., & Nahum-Shani, I. (2023). Predicting short-term suicidal thoughts in adolescents using machine learning: Developing decision tools to identify daily level risk after hospitalization. *Psychological Medicine*, 53(Guzmán Brand y Gélvez García, 2024), 2982-2991. <https://doi.org/10.1017/S0033291721005006>
- De Luca, G. P., Parghi, N., El Hayek, R., Bloch-Elkoubi, S., Peterkin, D., Wolfe, A., Rogers, M. L., & Galynker, I. (2024). Machine learning approach for the development of a crucial tool in suicide prevention: The Suicide Crisis Inventory-2 (SCI-2) short form. *PLOS ONE*, 19(Su et al., 2023), e0299048. <https://doi.org/10.1371/journal.pone.0299048>
- Gan, Y., Kuang, L., Xu, X. M., Ai, M., He, J. L., Wang, W., et al. (2025). Research on prediction model of adolescent suicide and self-injury behavior based on machine learning algorithm. *Frontiers in Psychiatry*, 15, 1521025. <https://doi.org/10.3389/fpsyt.2024.1521025>
- Gaynor, K., O'Reilly, M. D., Redmond, D., Nealon, C., Twomey, C., & Hennessy, E. (2023). A meta-analysis of targeted interventions for reducing suicide-related behaviour and ideation in adolescents: Implications for trial design. *Comprehensive Psychiatry*, 122, 152374. <https://doi.org/10.1016/j.comppsych.2023.152374>
- Guzmán Brand, V. A., & Gélvez García, L. E. (2024). Identificación de patrones a través de algoritmos de machine learning en los casos registrados de intentos suicidas en una ciudad de Colombia. *Psicoespacios*, 18(Gan et al., 2025), e1634. <https://doi.org/10.25057/21452776.1634>
- Haghighi, E. F., Czajkowski, N. O., & von Soest, T. (2023). Predicting suicide attempts among Norwegian adolescents without using suicide-related items: A machine learning approach. *Frontiers in Psychiatry*, 14, 1216791. <https://doi.org/10.3389/fpsyt.2023.1216791>
- Haghighi, E. F., Czajkowski, N., Walby, F. A., Qin, P., & Laeng, B. (2024). Suicide attempt risk predicts inconsistent self-reported suicide attempts: A machine learning approach using longitudinal data. *Journal of Affective Disorders*, 355, 495-504. <https://doi.org/10.1016/j.jad.2024.03.133>
- Haghighi, E. F., Laeng, B., & Czajkowski, N. (2023). Are false positives in suicide classification models a risk group? Evidence for "true alarms" in a population-representative longitudinal study of Norwegian adolescents. *Frontiers in Psychology*, 14, 1216483. <https://doi.org/10.3389/fpsyg.2023.1216483>
- Haghighi, E. F., Nes, R. B., Obaidi, M., Qin, P., Stänicke, L. I., Bekkhus, M., Laeng, B., & Czajkowski, N. (2024). Unveiling adolescent suicidality: Holistic analysis of protective and risk factors using multiple machine learning algorithms. *Journal of Youth and Adolescence*, 53(Stone et al., 2017), 507-525. <https://doi.org/10.1007/s10964-023-01892-6>
- Haroz, E. E., Kitchen, C., Nestadt, P. S., Wilcox, H. C., DeVlyder, J. E., & Kharrazi, H. (2021). Comparing the predictive value of screening to the use of electronic health record data for detecting future suicidal thoughts and behavior in an urban pediatric emergency department: A preliminary analysis. *Suicide and Life-Threatening Behavior*, 51(Smith et al., 2023), 1189-1202. <https://doi.org/10.1111/sltb.12800>
- Haroz, E. E., Rebman, P., Goklish, N., Garcia, M., Suttle, R., Maggio, D., Clattenburg, E., Mega, J., & Adams, R. (2024). Performance of machine learning suicide risk models in an American Indian population. *JAMA Network Open*, 7(Lee et al., 2024), e2439269. <https://doi.org/10.1001/jama-networkopen.2024.39269>
- Jankowsky, K., Steger, D., & Schroeders, U. (2024). Predicting lifetime suicide attempts in a community sample of adolescents using machine learning algorithms. *Assessment*, 31(Stone et al., 2017), 557-573. <https://doi.org/10.1177/10731911231167490>
- Kim, H., Son, Y., Lee, H., Kang, J., Hammoodi, A., Choi, Y., Kim, H. J., Lee, H., Fond, G., Boyer, L., Kwon, R., Woo, S., & Yon, D. K. (2024). Machine learning-based prediction of suicidal thinking in adolescents by derivation and validation in 3 independent worldwide cohorts: Algorithm development and validation study. *Journal of Medical Internet Research*, 26, e55913. <https://doi.org/10.2196/55913>
- Kim, K. W., Lim, J. S., Yang, C. M., Jang, S. H., & Lee, S. Y. (2021). Classification of adolescent psychiatric patients at high risk of suicide using the personality assessment inventory by machine learning. *Psychiatry Investigation*, 18(Zang et al., 2024), 1137-1143. <https://doi.org/10.30773/pi.2021.0191>
- Kusuma, K., Larsen, M., Quiroz, J. C., & Torok, M. (2024). Age-stratified predictions of suicide attempts using machine learning in middle and late adolescence. *Journal of Affective Disorders*, 365, 126-133. <https://doi.org/10.1016/j.jad.2024.08.043>
- Lee, H., Cho, J. K., Park, J., Lee, H., Fond, G., Boyer,

■ ANEXOS

Anexo 1

Checklist PRISMA Statement (ver pág. 15)

Section and Topic	Item	Checklist item	Location where is reported
TITLE			
Title	1	Identify the report as a systematic review.	
ABSTRACT			
Abstract	2	See the PRISMA 2020 for Abstracts checklist.	
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of existing knowledge.	
Objectives	4	Provide an explicit statement of the objective(s) or question(s) the review addresses.	
METHODS			
Eligibility criteria	5	Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses.	
Information sources	6	Specify all databases, registers, websites, organisations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted.	
Search strategy	7	Present the full search strategies for all databases, registers and websites, including any filters and limits used.	
Selection process	8	Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation tools used in the process.	
Data collection process	9	Specify the methods used to collect data from reports, including how many reviewers collected data from each report, whether they worked independently, any processes for obtaining or confirming data from study investigators, and if applicable, details of automation tools used in the process.	
Data items	10a	List and define all outcomes for which data were sought. Specify whether all results that were compatible with each outcome domain in each study were sought (e.g. for all measures, time points, analyses), and if not, the methods used to decide which results to collect.	
	10b	List and define all other variables for which data were sought (e.g. participant and intervention characteristics, funding sources). Describe any assumptions made about any missing or unclear information.	
Study risk of bias assessment	11	Specify the methods used to assess risk of bias in the included studies, including details of the tool(s) used, how many reviewers assessed each study and whether they worked independently, and if applicable, details of automation tools used in the process.	
Effect measures	12	Specify for each outcome the effect measure(s) (e.g. risk ratio, mean difference) used in the synthesis or presentation of results.	

Synthesis methods	13a	Describe the processes used to decide which studies were eligible for each synthesis (e.g. tabulating the study intervention characteristics and comparing against the planned groups for each synthesis (item #5)).	
	13b	Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data conversions.	
	13c	Describe any methods used to tabulate or visually display results of individual studies and syntheses.	
	13d	Describe any methods used to synthesize results and provide a rationale for the choice(s). If meta-analysis was performed, describe the model(s), method(s) to identify the presence and extent of statistical heterogeneity, and software package(s) used.	
	13e	Describe any methods used to explore possible causes of heterogeneity among study results (e.g. subgroup analysis, meta-regression).	
	13f	Describe any sensitivity analyses conducted to assess robustness of the synthesized results.	
Reporting bias assessment	14	Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases).	
Certainty assessment	15	Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome.	
RESULTS			
Study selection	16a	Describe the results of the search and selection process, from the number of records identified in the search to the number of studies included in the review, ideally using a flow diagram.	
	16b	Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded.	
Study characteristics	17	Cite each included study and present its characteristics.	
Risk of bias in studies	18	Present assessments of risk of bias for each included study.	
Results of individual studies	19	For all outcomes, present, for each study: (a) summary statistics for each group (where appropriate) and (b) an effect estimate and its precision (e.g. confidence/credible interval), ideally using structured tables or plots.	
Results of syntheses	20a	For each synthesis, briefly summarise the characteristics and risk of bias among contributing studies.	
	20b	Present results of all statistical syntheses conducted. If meta-analysis was done, present for each the summary estimate and its precision (e.g. confidence/credible interval) and measures of statistical heterogeneity. If comparing groups, describe the direction of the effect.	
	20c	Present results of all investigations of possible causes of heterogeneity among study results.	
	20d	Present results of all sensitivity analyses conducted to assess the robustness of the synthesized results.	
Reporting biases	21	Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed.	
Certainty of evidence	22	Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed.	

Anexo 2

Variable	Descripción
Referencia	Autores, año, título, revista.
País	Lugar donde se realizó el estudio.
Diseño del estudio	Retrospectivo, prospectivo, caso-control, cohorte, etc.
Objetivo	Propósito principal del estudio relacionado con predicción de suicidio/autolesión.
Tamaño muestral	Número total de participantes.
Edad media/rango de edad	Edad de los adolescentes incluidos.
Tipo de conducta analizada	Suicidio consumado, intento de suicidio, ideación suicida, autolesiones no suicidas.
Fuente de datos	EHR (Electronic Health Records), encuestas, autoinformes, registros hospitalarios, etc.
Tipo de modelo ML/IA usado	Random Forest, Support Vector Machine, Neural Networks, XGBoost, Deep Learning, otros.
Variables predictoras	Factores o variables incluidas en el modelo (ej.: síntomas depresivos, historia familiar, eventos vitales).
Métricas de desempeño	AUC (ROC), sensibilidad, especificidad, precisión, recall, F1 score, etc.
Principales resultados	Principales hallazgos sobre eficacia predictiva.
Factores de riesgo destacados	Variables predictoras con mayor peso/impacto.
Limitaciones del estudio	Limitaciones metodológicas reconocidas por los autores o detectadas.
Evaluación de calidad	Valoración de la calidad metodológica (usar herramienta PROBAST o similar).

Anexo 3

Evaluación de la calidad metodológica mediante el PROBAST

Con la finalidad de valorar sistemáticamente el riesgo de sesgo y la aplicabilidad de los estudios incluidos en esta revisión, se utilizó de forma retrospectiva el PROBAST (Prediction model Risk Of Bias Assessment Tool), herramienta destinada a evaluar estudios y validación de modelos predictivos.

Se ha centrado en cuatro dominios: participantes, predictores, resultado y análisis estadístico, estableciendo para cada uno una calificación de riesgo (bajo, moderado o alto). A partir de estas valoraciones, se ha determinado un juicio global de riesgo de sesgo y una estimación de la aplicabilidad de los estudios a contextos clínicos o preventivos reales.

Nº bibliografía	Referencia	Participantes	Predictores	Resultado	Análisis	Riesgo global de sesgo	Aplicabilidad
36	Haghish EF, Laeng B, Czajkowski N (2023). Are false	Bajo	Moderado	Alto	Moderado	Alto	Adecuada
24	Nes R, Obaidi M, Qin P, Stänicka LL, Haroz EE, Kitchen C, Nestadt PS, Wilcox HC, Cohen J, Wright-Berryman J, Rohlf L, Wright D, Czyz EK, Koo HJ, Al-Dajani N, King CA, Nahum-Shani I, Gan Y, Kuang L, Xu X-M, Ai M, He J-L, Wang W, Hono S, Haghish EF, Czajkowski NO, von Smet T (2023).	Bajo	Moderado	Alto	Moderado	Alto	Alta
38	Haghish EF, Czajkowski N, Walhov FA, Qin P, Xiling Yin, Dan Ma, Kejing Zhu, Deyun Li (2023).	Moderado	Bajo	Moderado	Moderado	Moderado	Limitada
41	Berryman J, Rohlf L, Wright D, Czyz EK, Koo HJ, Al-Dajani N, King CA, Nahum-Shani I, Gan Y, Kuang L, Xu X-M, Ai M, He J-L, Wang W, Hono S, Haghish EF, Czajkowski NO, von Smet T (2023).	Moderado	Moderado	Moderado	Moderado	Moderado	Alta
25	Haghish EF, Czajkowski N, Walhov FA, Qin P, Xiling Yin, Dan Ma, Kejing Zhu, Deyun Li (2023).	Bajo	Moderado	Alto	Moderado	Alto	Alta
32	X-M, Ai M, He J-L, Wang W, Hono S, Haghish EF, Czajkowski NO, von Smet T (2023).	Bajo	Moderado	Alto	Moderado	Alto	Adecuada
34	Czajkowski NO, von Smet T (2023).	Bajo	Moderado	Alto	Moderado	Alto	Adecuada
37	Czajkowski N, Walhov FA, Qin P, Xiling Yin, Dan Ma, Kejing Zhu, Deyun Li (2023).	Bajo	Moderado	Alto	Moderado	Alto	Alta
45	Kejing Zhu, Deyun Li (2023).	Bajo	Moderado	Moderado	Moderado	Moderado	Alta
33	Yinghao Chen, Gennou Yan, Zheng Orion Weller, Luke Sagers, Carl Hanson, Michael C.L. van Vuuren, K. van Mens, D. de Beurs, J. Raymond Su, James Rufus John, Pina-I Lin (2023).	Bajo	Moderado	Moderado	Moderado	Moderado	Adecuada
44	Sagers, Carl Hanson, Michael C.L. van Vuuren, K. van Mens, D. de Beurs, J. Raymond Su, James Rufus John, Pina-I Lin (2023).	Bajo	Moderado	Moderado	Moderado	Moderado	Alta
35	Hanson, Michael C.L. van Vuuren, K. van Mens, D. de Beurs, J. Raymond Su, James Rufus John, Pina-I Lin (2023).	Moderado	Moderado	Moderado	Moderado	Moderado	Adecuada
5	James Rufus John, Pina-I Lin (2023).	Bajo	Moderado	Alto	Moderado	Alto	Adecuada
42	Aseltine, Riddhi Doshi, Kun Chen, Michaela Servi, Silvia Chiaro, Elisa Musaj, Giovanni Moa Ponten, Oskar Flygare, Martin Ballander, Moa Chengxi Zang, Yu Hou, Daoming Lyu, Jun Jin, Shane Lim JS, Yang CM, Baek JW, Lee SY, Kim BN (2022).	Bajo	Bajo	Moderado	Moderado	Moderado	Alta
12	Silvia Chiaro, Elisa Musaj, Giovanni Moa Ponten, Oskar Flygare, Martin Ballander, Moa Chengxi Zang, Yu Hou, Daoming Lyu, Jun Jin, Shane Lim JS, Yang CM, Baek JW, Lee SY, Kim BN (2022).	Bajo	Bajo	Moderado	Moderado	Moderado	Adecuada
46	Flygare, Martin Ballander, Moa Chengxi Zang, Yu Hou, Daoming Lyu, Jun Jin, Shane Lim JS, Yang CM, Baek JW, Lee SY, Kim BN (2022).	Bajo	Moderado	Alto	Moderado	Alto	Adecuada
11	Hou, Daoming Lyu, Jun Jin, Shane Lim JS, Yang CM, Baek JW, Lee SY, Kim BN (2022).	Moderado	Bajo	Moderado	Moderado	Moderado	Alta
39	Baek JW, Lee SY, Kim BN (2022).	Bajo	Moderado	Moderado	Moderado	Moderado	Adecuada
10	Park J, Lee H, Fond G, Bover L, Kim H.J, Kim KW, Lim JS, Yang CM, Jang SH, Lee SY (2021).	Bajo	Moderado	Moderado	Moderado	Moderado	Alta
40	Yang CM, Jang SH, Lee SY (2021).	Bajo	Moderado	Moderado	Moderado	Moderado	Adecuada
13	H, Kang J, Hammongli A, Choi Jankowsky K, Steger D, Schwangers U, Haroz EE, Reisman P, Goklish N, Garcia M, Suttle R, Mo L, Li H, Zhu T (2022). Exploring the Suicide	Bajo	Moderado	Moderado	Moderado	Moderado	Alta
27	Steger D, Schwangers U, Haroz EE, Reisman P, Goklish N, Garcia M, Suttle R, Mo L, Li H, Zhu T (2022).	Bajo	Moderado	Moderado	Moderado	Moderado	Alta
20	P, Goklish N, Garcia M, Suttle R, Mo L, Li H, Zhu T (2022).	Moderado	Bajo	Moderado	Moderado	Moderado	Adecuada
43	Exploring the Suicide	Moderado	Moderado	Moderado	Moderado	Moderado	Adecuada

