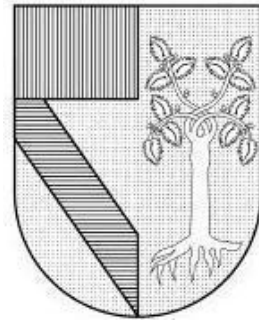


UNIVERSIDAD PANAMERICANA

**Facultad de Ciencias de la Salud
Escuela de Enfermería**



“Métodos de inteligencia artificial para la predicción de complicaciones en heridas: Revisión de alcance.”

TESINA

Q U E P R E S E N T A

**MARTÍNEZ AVILA NATALIA
SALGADO ARIZPE MARÍA VIRGINIA**

P A R A O B T E N E R E L G R A D O D E :

MAESTRÍA EN TERAPIA DE HERIDAS, ESTOMAS Y QUEMADURAS

DIRECTOR:

Dra. Laura Gabriela Chico Barba

CO-DIRECTOR:

Mtra. Blanca Estela López Hernández

Índice

1. Resumen.....	2
2. Marco teórico.....	3
3. Antecedentes.....	12
4. Planteamiento.....	14
5. Pregunta de investigación.....	15
6. Justificación.....	15
7. Objetivo de la revisión.....	16
8. Métodos.....	16
9. Resultados.....	19
10. Discusión y conclusiones.....	20
11. Referencias.....	23
Anexo 1: Estrategia de búsqueda en bases de datos.....	30
Anexo 2: Declaración de TRIPOD+AI.....	34
Anexo 3: Tablas de resumen heridas crónicas.....	36
Anexo 4: Tablas de resumen heridas quirúrgicas.....	45
Anexo 5: Tabla de evaluación a la adherencia metodológica con TRIPOD+AI heridas crónicas.....	56
Anexo 6: Tabla de evaluación a la adherencia metodológica con TRIPOD+AI con heridas quirúrgicas.....	57

Métodos de inteligencia artificial para la predicción de complicaciones en heridas: revisión de alcance

1. Resumen

Antecedentes: La Inteligencia Artificial (IA) es una herramienta innovadora con alto potencial en el ámbito clínico, especialmente en heridas quirúrgicas y crónicas. Su aplicación busca mejorar el diagnóstico, pronóstico y seguimiento de pacientes reduciendo la carga laboral. Sin embargo, su constante innovación requiere realizar una evaluación de dicha evidencia. El objetivo del estudio fue revisar la evidencia científica existente sobre el uso de IA en el diagnóstico, pronóstico y tratamiento de heridas crónicas: pie diabético, LPP, dehiscencia quirúrgica y úlceras vasculares.

Métodos: Se realizó una revisión de alcance mediante una búsqueda en PubMed y CINAHL. Empleando términos MeSH como: *chronic wound, IA, surgical site infection, wound dehiscence, wound healing, prognosis*. Incluyendo estudios publicados entre el año 2020 al 2025. Excluyendo artículos por idioma, estudios en animales o cadáveres y revisiones sistemáticas. Fueron seleccionados mediante plataforma Rayyan para estudios de revisión, partiendo de 772 más 8 artículos de una búsqueda manual, de los cuales 276 fueron eliminados por duplicidad. Se analizaron 44 artículos relevantes: 24 enfocados en heridas quirúrgicas y 20 en heridas crónicas.

Resultados: En heridas quirúrgicas, el algoritmo de IA identifica principalmente factores de riesgo para predecir infecciones, mientras que en las heridas crónicas se enfocan en el análisis de imágenes. La mayoría de los estudios son retrospectivos y unicéntricos, con bases de datos pequeñas, de muestra reducidos con riesgo de sobreestimación de datos y requirieron de métodos adicionales para calibrarlas. En cuanto a la adherencia a los criterios TRIPOD+AI en los artículos de heridas quirúrgicas, ninguno mostró una alta adherencia, 22 artículos presentaron adherencia moderada, mientras que en las heridas crónicas, sólo uno mostró una alta adherencia, y el resto fueron de adherencia moderada.

Conclusión: Los hallazgos de esta revisión de alcance evidencian la necesidad de estudios prospectivos, multicéntricos con validación externa. La IA puede fortalecer los sistemas de salud, reducir la carga laboral y mejorar la eficacia terapéutica.

2. Marco teórico

Definición de heridas

Una herida es definida por la WOCN (Wound, Ostomy and Continence Nurse society) como “la disrupción en la estructura y función normal de la piel y los tejidos blandos subyacentes y puede estar relacionada con una variedad de etiologías” (1). Estas son clasificadas en agudas o crónicas, con base en el autor estas pueden tener diferentes criterios en su definición.

La Doctora Elena Conde hace referencia a ambas heridas, dando a entender que las heridas agudas son aquellas que evolucionan en tiempo y forma esperada siguiendo el proceso de cicatrización (2). Con esto podremos entender que las heridas crónicas se estancan en el proceso de cicatrización teniendo una fase inflamatoria prolongada lo cual no permite su cierre. La WOCN define que una herida aguda es aquella que cicatriza en orden y tiempo sin requerir seguimiento a largo plazo mientras que las heridas crónicas además de considerar que se quedan estancadas en el proceso de cicatrización no cierran en un periodo de 4 semanas (1).

Heridas crónicas

Las heridas crónicas se clasifican en diferentes grupos según su etiología: úlceras venosas, úlceras arteriales, pie diabético y lesiones por presión. (3)

Las úlceras venosas son definidas por The American Venous Forum como “defecto de la piel de espesor total, con mayor frecuencia en la región del tobillo, que no cicatriza de forma espontánea y se mantiene debido a la enfermedad venosa crónica confirmada por un estudio de dúplex venoso”.(4) Este tipo de lesiones representan el 75-80% de todas las úlceras de extremidad inferior con una incidencia de 2-5 casos

nuevos por 1000 habitantes/año (3). Son clasificadas según el Sistema CEAP clínico, etiológico, anatómico y fisiopatológico.

Tabla 1. Clasificación CEAP. Componente clínico ©

Clase C	Descripción
C0	Sin signos visibles ni palpables de enfermedad venosa
C1	Telangiectasias o venas reticulares
C2	Venas varicosas
C3	Edema
C4	Cambios en la piel y el tejido subcutáneo secundarios a enfermedad venosa crónica
C4a	Pigmentación o eccema
C4b	Lipodermatoesclerosis o atrofia blanca
C4c	Corona flebectásica (dilataciones venosas en forma de corona)
C5	Úlcera venosa cicatrizada
C6	Úlcera venosa activa

Traducido y adaptado de: Sociedad Española de Angiología y Cirugía Vascul ar (SEACV). Guía práctica para el diagnóstico y tratamiento de la insuficiencia venosa crónica.

El estándar de oro para el manejo de estas lesiones es la terapia de compresión combinado con terapia adyuvantes para el cierre de la herida. (5)

Las úlceras arteriales están relacionadas con la enfermedad arterial periférica, también pueden ser conocidas como úlceras isquémicas. La principal etiología de estas lesiones se debe a la obstrucción aterosclerótica que desarrolla una isquemia crónica principalmente en zonas como el pie, tobillo o parte distal de la pierna en ellas es típica la necrosis y un lecho pálido (2,6).

La sintomatología que puede ser generada por estas úlceras son dolor profundo e intenso, claudicación que puede ser valorada con la escala de Fontaine o Rutherford. (7)

Tabla 2. Tabla comparativa de la Clasificación de Fontaine y Rutherford

FONTAINE		RUTHERFORD	
Estadio	Clínica	Estadio	Clínica
I	Asintomática	1	Asintomática
IIa	Claudicación intermitente no invalidante	2	Claudicación leve
IIb	Claudicación intermitente invalidante	3	Claudicación grave
III	Dolor en reposo	4	Dolor en reposo
IV	Úlceras o gangrena	5	Lesiones tróficas
		6	Gangrena

Tomado de: Fontaine R, Kim M, Kieny R. Surgical treatment of peripheral circulation disorders. *Helv Chir Acta.* 1954;21(5-6):499-533. y American College of Cardiology/American Heart Association. 2024 ACC/AHA/AACVPR/APMA/ABC/SCAI/SVM/SVN/SVS/SIR/VESS Guideline for the Management of Lower Extremity Peripheral Artery Disease. *Circulation.* 2024.

El tratamiento de las úlceras arteriales consta de mejorar la perfusión arterial al igual que controlar los factores de riesgo principalmente aquellos que son cardiovasculares, en cuanto a la herida se pueden proporcionar cuidados locales adecuados. En casos severos cuando la revascularización no es factible y se quiere evitar la sobreinfección se puede considerar la amputación de la extremidad. (6,7)

La diabetes mellitus en descontrol puede generar heridas como lo es el pie diabético siendo una de las principales causas de reducción de calidad de vida así como de costos económicos para el paciente y su familia. Dentro de los principales factores que influyen a la aparición de una úlcera se encuentran enfermedad arterial periférica, neuropatía periférica y un componente mixto. (8)

Actualmente existen diferentes clasificaciones que nos permite estadificar de acuerdo a la gravedad de la lesión como lo es SINBAD. La guía de práctica clínica

de International Working Group on the Diabetic Foot (IWGDF) sugiere su uso para su clasificación simple y lograr su reproducibilidad en diferentes entornos clínicos. Esta escala valora Sitio, Isquemia, Infección bacteriana, Área y Profundidad. (8)

Tabla 3. Clasificación de SINBAD

Categoría	Definición	Puntuación
Localización	Antepié	0
	Mediopié y retropié	1
Isquemia	Flujo sanguíneo intacto del pie: al menos un pulso palpable	0
	Evidencia clínica de flujo sanguíneo reducido del pie	1
Neuropatía	Sensación protectora intacta	0
	Pérdida sensación protectora	1
Infección bacteriana	Ausente	0
	Presente	1
Área úlcera	Úlcera < 1 cm²	0
	Úlcera ≥1 cm²	1
Profundidad	Úlcera limitada a la piel y tejido subcutáneo	0
	Úlcera que alcanza el músculo, tendón o más profunda	1
Puntuación total posible		0 - 6

Tomado de: Monteiro-Soares M, Hamilton EJ, Russell DA, Srisawasdi G, Boyko EJ, Mills JL, Jeffcoate W, Game F, et al. Guidelines on the classification of foot ulcers in people with diabetes (IWGDF 2023 update). *Diabetes Metab Res Rev.* 2024;40(3):e3648. doi:10.1002/dmrr.3648

La IWGDF sugiere que el pilar del tratamiento es el “offloading” liberar la presión del pie para que este pueda cicatrizar de forma adecuada utilizando dispositivos no removibles de primera elección. En cuanto al cuidado de la herida es importante tener una inspección por el personal capacitado, desbridamiento quirúrgico,

mantener un ambiente húmedo, cuidado la piel perilesional y el manejo y control de la infección. Por otro lado la guía nos dice que es importante individualizar cada caso ya que en algunos casos se podrá sugerir la revascularización o el uso de terapias avanzadas. (8)

Para fines de este trabajo nos enfocaremos en las lesiones ocasionadas por presión o cizallamiento. El Grupo Nacional para el Estudio y Asesoramiento en Úlceras por Presión y Heridas Crónicas (GNEAUPP) las define como “Una lesión localizada en la piel y/o el tejido subyacente por lo general sobre una prominencia ósea, como resultado de la presión, o la presión en combinación con las fuerzas de cizalla...” (9)

Como se ha mencionado antes la principal etiología es la constante presión que a su vez causa una hipoperfusión de los tejidos, las lesiones son dependientes del tiempo de permanencia de la presión. Pudiéndose clasificar en diferentes estadios, como lo hace GNEAUPP (9)

El tratamiento de este tipo de lesiones dependerá del estadio en el que se encuentre y el porcentaje de tejidos en la herida. Sin embargo, el pilar fundamental es eliminar la presión y disminuir al máximo las fuerzas de cizallamiento y fricción (9), es importante valorar e individualizar a cada uno de los pacientes.

Heridas quirúrgicas

Por otro lado existen las complicaciones de heridas quirúrgicas que consisten en la interrupción de la curación de una herida por incisión normal después de una cirugía incluyendo infección de sitio quirúrgico (ISQ), dehiscencia quirúrgica (DHQ), hipergranulación, maceración de la piel perilesional, cicatrización anómala, daño por adhesivos sanitarios (MARSÍ). En este trabajo sólo profundizaremos en las dehiscencias quirúrgicas e infecciones de sitio quirúrgico. (10).

En base a las recomendaciones que describe International Surgical Wound Complications Advisory panel (ISWCAP). Considera que la ISQ puede suceder a

partir de 30 días después de un procedimiento quirúrgico sin implantes o hasta un año cuando se usan implantes, y este se clasifica de la siguiente forma: (10)(11)

Tabla 4: Definición de Infección de Sitio Quirúrgico (ISQ) según Horan et al., 1992

Definición de ISQ
Puede ser superficial, profunda o de órgano espacio. Se produce hasta 30 días después de la intervención quirúrgica e implica uno o todos los elementos siguientes:
1. Drenaje purulento: Con o sin confirmación de laboratorio del lugar de la incisión.
2. Organismo aislado: Cultivo positivo de líquido o tejido de la incisión obtenido asépticamente.
3. Signos o síntomas de infección: Dolor, sensibilidad, hinchazón localizada, enrojecimiento o calor, o apertura deliberada de la incisión por parte del cirujano (incluso con cultivo negativo).
4. Diagnóstico clínico: Diagnóstico de ISQ por parte del cirujano o el médico responsable del tratamiento.

Tomado de :Horan TC, Gaynes RP, Martone WJ, Jarvis WR, Emori TG. CDC definitions of nosocomial surgical site infections, 1992: a modification of CDC definitions of surgical wound infections. Infect Control Hosp Epidemiol. 1992;13(10):606–8. doi:10.2307/30148488

La DHQ es definida como la separación de los márgenes de una incisión quirúrgica cerrada, con o sin exposición o protrusión de tejidos, órganos o implantes subyacentes. Puede o no estar asociada a infección y puede afectar una o varias capas del tejido. La clasificación de WUHUNS 2018 propone:

Tabla 5: Clasificación de Dehiscencia de Herida Quirúrgica de la WUHUNS 2018

Grado de DHQ	Descripción
1	Solo la epidermis; tejido subcutáneo no visible Sin signos o síntomas clínicos de infección
1a	DHQ de grado 1, más signos y síntomas clínicos de infección
2	Capa subcutánea expuesta, fascia no visible Sin signos o síntomas clínicos de infección
2a	DHQ de grado 2, más signos y síntomas clínicos de infección
3	Capas subcutáneas y fascia expuestas Sin signos o síntomas clínicos de infección
3a	DHQ de grado 3, más signos y síntomas clínicos de infección
4	Cualquier zona de dehiscencia fascial con órganos o espacios, vísceras, implantes o hueso expuestos Sin signos o síntomas clínicos de infección
4a	DHQ de grado 4, más signos y síntomas clínicos de infección

Tomado de: Zhang J, Li Q, Wang Y, et al. Clinical classification of surgical wound dehiscence: WUHUNS 2018 consensus. *J Surg Wound Care*. 2018;7(3):112–118.

Las heridas quirúrgicas se pueden abordar de diferentes formas, tomando en cuenta que siempre es importante individualizar cada caso de acuerdo al grado de severidad. En estas se puede incluir el uso de terapias avanzadas como lo es la terapia de presión negativa y en caso de infección considerar el uso de antibioticoterapia sin abusar de estos.

Inteligencia Artificial en Salud

En los últimos años la inteligencia artificial (IA) se ha convertido en una herramienta que puede transformar la medicina. Diferentes artículos han dado definiciones para poder explicar el concepto de la IA y sus divisiones.

La IA puede definirse como una herramienta que automatiza y simula tareas intelectuales humanas como el análisis de datos, la cual incluye el Machine Learning (ML) y el Deep Learning (DL) (11). Por otro lado, International Business Machine (IBM) define la IA como “una tecnología que permite a las computadoras y máquinas simular el aprendizaje humano, la comprensión, la resolución de problemas, la toma de decisiones, la creatividad y la autonomía”.(12)

A lo largo de los años la IA ha ido evolucionando englobando diferentes conceptos como lo son el ML, DL y las redes neuronales. El ML es un sistema de algoritmos que identifica patrones de datos para hacer sus predicciones, mejorando a través de la experiencia sin necesidad de una programación explícita (13). Algunos ejemplos de modelos son regresión lineal, árboles de decisión, random forest, support vector machine (SVM) entre otros.

La evolución del machine learning (ML) deriva en el deep learning (DL) lo que lo hace más específico y rápido. El DL cuenta con una base de datos extensa en la cual existen diferentes algoritmos que lo ayudan a procesar la información como lo haría un cerebro humano, es por esto que no requiere una intervención adicional. Esto sucede a partir de un procesamiento de datos mediante unidades de computación llamadas “neuronas”, las cuales están organizadas en secciones ordenadas a las cuales se les denomina “capas”, es decir que la técnica de aprendizaje en el DL son las redes neuronales(14).

Las redes neuronales (ANN) pueden ser consideradas como la unidad básica de la IA. Su funcionamiento es basado en algoritmos de aprendizaje inspirados por la biología neuronal. (13) Para entender el funcionamiento de las ANN es importante recordar que es análogo a una neurona. Están conformadas por nodos que

contienen datos por los cuales se comunican por estructuras parecidas a los axones y dendritas (11,12).

Cada IA ha evolucionado dependiendo de su fin, algunas utilizan el procesamiento de imágenes para discriminar características. Esto es útil en el campo de la medicina ya que ayuda al procesamiento en grandes cantidades de datos visuales para realizar diagnósticos y decisiones terapéuticas de forma oportuna impactando de forma positiva a este campo (13).

Por otro lado existen inteligencias artificiales que su técnica se basa en el análisis de datos generando escenarios posibles en los cuales se utilizan los árboles de decisiones (decision tree, random forest, XG Boost). Podríamos entender que estas hacen instrumentos de valoración mediante factores predictivos dando datos favorables para la estimación de riesgos. Es por eso que tienen utilidad para identificar pacientes con probabilidad de complicaciones. Los últimos años ha sido implementado en ambientes quirúrgicos por su eficacia, disminuyendo la morbi-mortalidad en complicaciones quirúrgicas, realizando intervenciones oportunas. (15,16).

IA en heridas

Para la evaluación de heridas es crucial la toma de fotografías, ya que nos permite dar un seguimiento. No solo nos habla del progreso, incluso nos ayuda a revalorar nuestros tratamientos. Es por esto que se ha implementado su uso en las IA, la digitalización de fotografías ha permitido el entrenamiento de estas y por lo tanto la creación de algoritmos que permiten la estratificación, estadificación, diagnóstico y posible evolución de las heridas crónicas. Por ejemplo DFU Care para pacientes con úlceras de pie diabético (17), YOLO (you only live once) para lesiones por presión entre otras (18).

3. Antecedentes

La IA es una disciplina científica que nació oficialmente en 1956 en el Dartmouth College, en Hanover (Estados Unidos). Desde entonces ha sufrido diferentes transformaciones que le han permitido ser aplicada como una herramienta en diferentes campos (19). En el mundo de las heridas ha sido útil para la detección, localización, clasificación, evaluación, diagnóstico y monitoreo. Por ejemplo:

1. Estudio molecular en heridas: identifica a SCUBE1 y RNF103-CHMP3 como objetivos terapéuticos para las úlceras del pie diabético (UPD) posibles biomarcadores de la UPD supone un avance significativo en la comprensión de la base molecular de la enfermedad. (20)
2. Detección, localización y clasificación: Modelo SNN con técnicas de explicabilidad para mejorar la detección de úlceras de pie diabético. (21)
3. Predicción de infecciones: Se utilizó un modelo de SENIC -LR y Naive para mejorar el desempeño predictivo para SSI para predecir infecciones de sitio quirúrgico en cirugías abdominales (15).
4. Medición en las heridas: Uso de aplicaciones que utilizan aprendizaje automático como herramientas de procesamiento de imágenes para la evaluación y monitorización inteligente de heridas (22).

En heridas crónicas la literatura está enfocada en realizar modelos de procesamiento de imagen que permiten la medición y/o de factores predictivos de riesgo. El uso de imágenes es útil porque existen limitaciones al haber diferentes tipos de heridas, además de que las presentaciones son variadas. El entrenamiento de estas Inteligencias Artificiales fue realizado por medio de imágenes localizadas en bases de datos obtenidas por universidades, hospitales y clínicas.

Sin embargo, entre las limitantes que enfrentan los estudios están la falta de diversidad en las imágenes (heridas, grupos étnicos), falta de validación en entornos clínicos, tamaño de muestra reducido para el entrenamiento de IA y alto costo computacional. Es importante destacar que en cuanto a las imágenes la eficacia

depende de la calidad de los datos y es necesario unificar los criterios de la toma de fotografías.(17,23–25)

En 2022 se realizó una revisión narrativa en pacientes con UPD, que describe el rol que ha desempeñado la IA en este ámbito. Uno de los aspectos a destacar mencionados es la monitorización remota de estas heridas ya que empodera a los pacientes y profesionales para mejorar la atención, aunque aún siguen en desarrollo estos algoritmos digitales para la mejora de su sensibilidad y especificidad. (26)

Además del uso del análisis mediante imágenes también se pueden emplear algoritmos específicos. En la actualidad estos sirven en expedientes clínicos para identificar a sujetos con antecedentes familiares de enfermedades hereditarias o un mayor riesgo de padecer enfermedades crónicas. (27)

En heridas quirúrgicas es una herramienta útil para identificar sujetos con mayor riesgo de desarrollar complicaciones. (15) En 2024 se publicó una revisión sistemática donde se incluyeron modelos de ML que predecían ISQ en pacientes adultos post operados donde se concluyó que aún se requiere calibración de los modelos y validación externa de los mismos. (26)

A pesar de que la IA lleva años en desarrollo aún existen áreas de oportunidad en cuanto al aspecto ético y la privacidad de los datos. (25)

4. Planteamiento

Las heridas crónicas y quirúrgicas representan un desafío constante, debido a su elevada morbilidad por costos asociados a su manejo, tiempos de cicatrización y su impacto en la calidad de vida. Las diferentes heridas como úlceras por presión, pie diabético, úlceras venosas y las complicaciones de heridas postquirúrgicas como infecciones o dehiscencias representan no solo un impacto en la calidad de vida de los pacientes y sus familiares sino también una carga en los sistemas de salud a nivel global.

Debido a la gran demanda, en ocasiones los métodos tradicionales de diagnóstico, evaluación y seguimiento pueden resultar insuficientes para la identificación oportuna de complicaciones, aumentando el riesgo de una mala evolución.

Con la evidencia antes expuesta podemos concluir que la Inteligencia Artificial está en un desarrollo constante, generando información continuamente en los últimos cinco años, posicionándose como una herramienta innovadora y de gran utilidad en el área de heridas. De esta forma nos ayuda a predecir eventos adversos con mayor precisión en comparación a los métodos tradicionales. (15,28)

Es por esto que se considera realizar un estudio que analice los artículos científicos publicados entre los años 2020 a 2025, con el objetivo de identificar los avances relevantes del uso de IA aplicada al manejo de heridas así como proponer estrategias que permitan su implementación efectiva y generalizada a los diferentes ámbitos clínicos.

5. Pregunta de investigación

Por lo anterior, la pregunta de investigación es: ¿Qué evidencia existe sobre el uso de inteligencia artificial para la evaluación, diagnóstico, clasificación y predicción de complicaciones en pacientes con heridas crónicas y quirúrgicas?

6. Justificación

Las heridas crónicas representan una prevalencia global del 1% al 2% con mayores tasas en adultos mayores. Un artículo elaborado en Estados Unidos reporta que las heridas crónicas afectan al 2.5% de su población adulta principalmente en pacientes geriátricos teniendo un impacto en la calidad de vida y costos de atención médica elevados. (29,30)

Por otro lado, un estudio realizado en Tabasco analiza que la prevalencia de infección de sitio quirúrgico reportado es de un 15.8% solo en las cirugías de colecistectomía. (31) En México se estima que la prevalencia de úlceras es de un

10.1% siendo más alta en personas de 71-80 años. Cabe señalar que las unidades médicas incluidas en el presente estudio representan menos del 1% de la cobertura actual del Sistema Nacional de Salud en el país. En este estudio se estimaron los costos directos del tratamiento siendo de \$46,563,070.76 (ambulatorio) y \$1,864,124,436.89 (hospitalización). (32)

Por lo anteriormente expuesto, la ausencia de un adecuado control epidemiológico por parte del Sistema Nacional de Vigilancia Epidemiológica sobre heridas dificulta una gestión clínica eficiente. Es por eso que los métodos de inteligencia artificial podrían ser una herramienta valiosa que ayude a optimizar el uso de personal y recursos, mejorando así el tratamiento y prevención de complicaciones.

Debido a esta problemática, surge la necesidad de recopilar y sistematizar la evidencia científica ya existente sobre el uso de IA para la evaluación, diagnóstico, clasificación y predicción de complicaciones en pacientes con heridas crónicas y quirúrgicas, a fin de identificar su utilidad clínica, limitaciones y áreas de oportunidad para la aplicación en centros especializados en heridas.

7. Objetivo de la revisión

Objetivo General

Evaluar la evidencia existente de los de los sistemas de inteligencia artificial en la predicción de complicaciones en las heridas

Objetivos Específicos

- Analizar la evidencia sobre el uso de herramientas basadas en inteligencia artificial para la predicción y seguimiento de heridas quirúrgicas
- Evaluar la evidencia científica sobre la aplicación de inteligencia artificial en el diagnóstico, pronóstico y tratamiento de heridas crónicas
- Determinar la calidad científica de los artículos incluidos en la revisión

8. Métodos

a. Fuentes de información

Se utilizaron las base de datos bibliográficas como PubMed y CINAHL

b. Criterios de inclusión, exclusión de los estudios

Criterios de Inclusión

- Artículos publicados entre el enero del 2020 a junio del 2025
- Artículos donde se haga uso de cualquier método de IA para la evaluación de heridas crónicas y quirúrgicas
- Artículos donde se realice evaluación de heridas crónicas y/o quirúrgicas mediante el uso de imágenes y algoritmos.

Criterios de Exclusión

- Artículos que no se encuentren indexados o que estén en preprint (en prensa)
- Artículos publicados en idiomas diferentes al español o el inglés.
- Estudios realizados en animales o cadáveres
- Revisiones sistemáticas.

Criterios de Eliminación

- Artículos que hayan sido withdrawn, retracted o removed de las base de datos

c. Estrategia de búsqueda

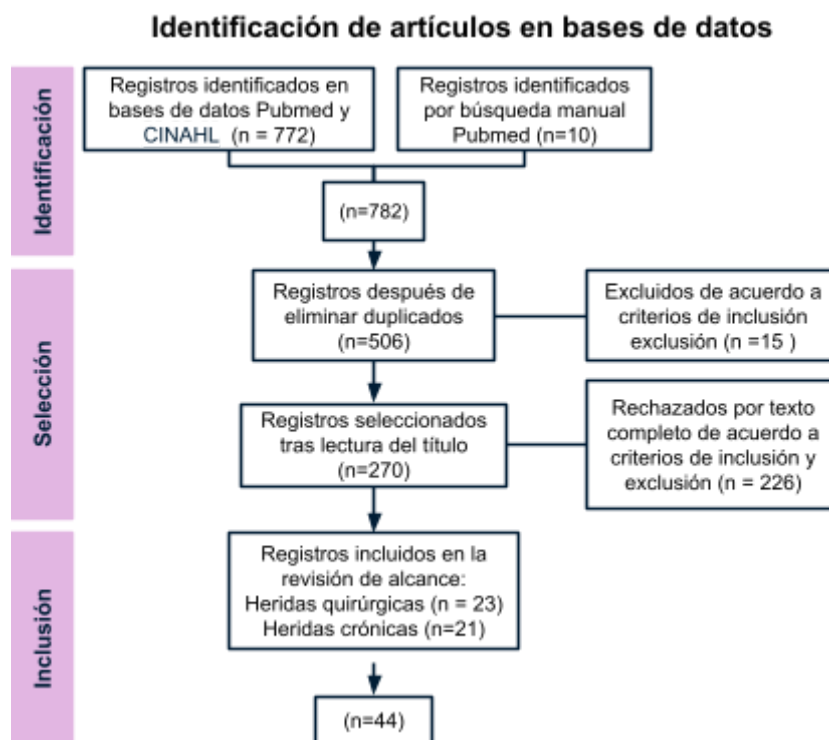
En general, se utilizaron los términos MeSH chronic wound, Machine Learning, Diabetic foot Ulcer, pressure ulcer, artificial intelligence, wound infection, deep learning, surgical site infection, wound healing , prognosis y se ajustaron para realizar diferentes búsquedas, como se muestran en el anexo #1.

Se adjunta la tabla de la estrategia de búsqueda en el Anexo #1.

d. Descripción de variables a revisar en los artículos

Para la revisión de artículos se consideraron como variables principales aquellas que estuvieran relacionadas con la Inteligencia Artificial (IA), incluyendo términos como machine learning (ML), deep learning (DL) entre otros. Los artículos también deben abordar algún enfoque clínico como: diagnóstico, clasificación, prevención o pronóstico en el contexto de heridas crónicas o quirúrgicas.

e. Diagrama de flujo de selección de estudios (PRISMA)



f. Aspectos éticos

Este estudio es una revisión de alcance, en el cual no es necesaria la participación directa de personas al no recopilar datos personales o sensibles, es por esto que no se requiere una aprobación del comité de ética universitario. Este trabajo se basa en el análisis de artículos científicos previamente publicados.

g. Descripción de la presentación de la evidencia

Se elaboraron tablas de extracción de datos para sistematizar la información de cada estudio incluido. Dichas tablas se enfocan en variables como: título, año, país, población y tipo de herida, tipo de estudio, programa de IA utilizado, propósito de IA, resultados, limitaciones y conclusiones.

Posteriormente se llevó a cabo una evaluación de la adherencia que presentaban los artículos al checklist de TRIPOD+AI. La cual nos permitió valorar la fiabilidad de cada estudio.

h. Evaluación de la adherencia metodológica

Para valorar el cumplimiento sobre el reporte de informes que hacen uso de modelos de predicción se utilizó la declaración de TRIPOD+AI encontrado en el anexo 2 que ofrece una guía para calificar los informes de estudios enfocados en modelos de predicción con modelos de regresión o aprendizaje automático. (33)

Esta herramienta propone una lista de verificación de 27 ítems principales que permiten calificar los diferentes componentes de un artículo. Estos se encuentran distribuidos de la siguiente forma: título (ítem 1), el resumen (ítem 2), la introducción (ítem 3 y 4), los métodos (ítem 5 al 17), las prácticas de ciencia abierta (ítem 18), la participación del paciente y del público (ítem 19), los resultados (ítem 20-24) y la discusión (del ítem 25 al 27). Además algunos están compuestos por subítems, totalizando en una checklist de 52 elementos. (33)

Con base en los criterios TRIPOD+AI, tanto en los artículos de heridas quirúrgicas como de heridas crónicas la mayoría de los artículos presentó una adherencia moderada, demostrando que existe la necesidad de realizar difusión de los criterios de evaluación y la adherencia a estos para mejorar la calidad metodológica para investigaciones futuras.

9. Resultados

Los artículos fueron seleccionados mediante la plataforma Rayyan para estudios de revisión, partiendo de 772 y siendo agregados 10 artículos más provenientes de una búsqueda manual, de los cuales 276 fueron eliminados por duplicidad. Se analizaron 44 artículos relevantes, 23 enfocados en heridas quirúrgicas y 21 en heridas crónicas. (15–18,20,21,28,34–70)

En heridas quirúrgicas, los modelos de inteligencia artificial se orientaron principalmente a la identificación de factores de riesgo y predicción de infecciones en sitio quirúrgico, incluyendo variables clínicas, quirúrgicas y de laboratorio para su modelo. Por el contrario, las heridas crónicas en su mayoría se centraron en el

análisis de imágenes digitales, utilizando técnicas de aprendizaje profundo para la clasificación de estadios, identificando bordes y estimación de tejidos.

El tipo de herida crónica principalmente evaluado fueron las úlceras de pie diabético (n=12) seguidas por las lesiones por presión (n=7) y heridas crónicas de diferente etiología (n=2). La mayoría de los estudios incluidos fueron realizados en el continente asiático, destacando India (n=4), China (n=4) y Estados Unidos (n=3) como los países con mayor número de publicaciones. El año con mayor número de estudios fue el 2024 (n=8). La mayor parte de los estudios son retrospectivos (n=18) y unicéntricos (n=19). Solamente 3 de las investigaciones realizaron validación cruzada y solo 9 artículos fueron probados en entornos clínicos.

Por otro lado, en las heridas quirúrgicas destacan las cirugías ortopédicas (n=8) y cirugías abdominales/colorrectales (n=6). Al igual que en las heridas crónicas el año con mayor número de publicaciones fue el 2024 (n=7), con predominio en Estados Unidos (n=8) y China (n=6), seguidos de países europeos (n=5), Canadá, Corea del Sur, Tailandia y Ruanda. En cuanto al diseño metodológico, fueron cohortes retrospectivas (n=21) frente a sólo dos estudios prospectivos. En general fueron unicéntricos (n=13) sin validación externa. Finalmente, la mayoría no reporta pruebas en entornos clínicos reales (n=18).

Ambos grupos muestran que el tamaño de muestra es reducido en las bases de datos, incrementado el riesgo de sobreajuste y sobreestimación de los resultados. Es por ello que varios trabajos requieren métodos adicionales de calibración y validación cruzada para mejorar la fuerza de sus resultados. En cuanto a las heridas quirúrgicas, no se identificó un reporte del tiempo de cicatrización, lo cual causó una limitante metodológica ya que no se pudo establecer si estas podrían ser clasificadas como heridas crónicas limitando los modelos propuestos.

En cuanto a la adherencia a los criterios TRIPOD+AI, en el caso de las heridas crónicas los estudios alcanzaron en su mayoría un nivel moderado teniendo una mediana del 61.5% y un p25 de 55.8% y p75 de 63.5%. Todos los trabajos fueron

ubicados en la categoría de calidad moderada (50-69%) sin que ninguno lograra tener una alta adherencia (>70%).

Por otro lado, los estudios de heridas quirúrgicas mostraron una mayor dispersión en sus resultados. La mediana de adherencia fue de 55.8% con un rango de p25 51.9% y p75 65.4%. La mayoría de los artículos (n=22) se ubicaron en una categoría de calidad moderada, aunque hubo uno de ellos que presentó baja adherencia (<50%) lo que muestra heterogeneidad.

Las tablas de resumen se muestran en los Anexos #3 y #4, así como la evaluación de adherencia metodológica mediante TRIPOD +AI en Anexos #5 y #6

10. Discusión y conclusiones

La aplicación de la Inteligencia Artificial (IA) en el manejo de heridas crónicas y quirúrgicas se ha convertido en una herramienta innovadora con alto potencial para mejorar la práctica clínica (71).

En la revisión sistemática de Riefs y Jiménez del 2025 (24) analizaron múltiples investigaciones centradas en heridas crónicas, principalmente lesiones por presión y úlceras por pie diabético. En su estudio muestran que hubo un desempeño sobresaliente en modelos de deep learning, redes convencionales U-Net y EfficientNetB0, ya que podían realizar una segmentación de tejidos y clasificación de las lesiones en los diferentes estadios. Al igual que en esta revisión, la mayoría de los artículos utilizaron imágenes de heridas para aplicar algoritmos basados en técnicas de DL para el análisis de las imágenes y además encontraron artículos que aplican el uso de algoritmos de ML para el procesamiento de datos del registro clínico de los pacientes. (24)

En una revisión sistemática del 2024 que evaluaba la calidad metodológica de modelos de aprendizaje automático para la predicción de infecciones del sitio quirúrgico postoperatorio concluyó que la mayoría de los artículos recopilados no

contaban con una validación externa y que solo 2 de los 24 estudios incluidos reportaron métodos de calibración (26). Por otro lado, el trabajo de Tabja Bortesi y colaboradores observó que a pesar de que sus estudios mostraron resultados prometedores (valores AUC de 0.65 a 1.0), la mayoría de los estudios carecía de validación externa y presentaban alto riesgo de sesgo por falta de estandarización en la captura de imágenes y en la definición de los criterios de infección. (72)

En nuestro análisis, los resultados concuerdan parcialmente con los hallazgos de dichas revisiones. Coincidimos en que el uso de IA en el manejo de heridas quirúrgicas aún se encuentra en etapas iniciales y requiere protocolos estandarizados que garanticen la calidad de imágenes y la consistencia diagnóstica. Sin embargo, nuestros resultados sugieren que hay avances en la aplicación de modelos más robustos y en la inclusión de variables complementarias, ya que ayudan a la precisión diagnóstica.

Su integración busca optimizar la toma de decisiones y reducir la subjetividad en la evaluación de las heridas. Existen diferentes estudios alrededor del mundo que han demostrado que el uso de IA contribuye a la identificación más precisa de factores de riesgo, predicción de complicaciones e incluso estimación en el tejido.

Entre las limitaciones se reconoce que aún no existe un consenso general y universal sobre herramientas utilizadas para el reporte. Los artículos revisados presentaban heterogeneidad, donde se caracterizaban por escasa validación externa, estudios unicéntricos, retrospectivos y diseños metodológicos similares lo que limita la comparabilidad. Además, la rápida evolución de la tecnología podría requerir el desarrollo de herramientas especializadas generando cambios en los criterios actuales.

Una de las fortalezas de nuestro estudio es la aplicación del instrumento TRIPOD+AI, que permite evaluar de forma estructurada la adherencia metodológica en los estudios incluidos. Esta revisión aporta una visión comparativa de las tendencias sobre el uso de IA en el cuidado de heridas, pudiendo identificar que el

análisis de imágenes es más común en heridas crónicas utilizando las redes neuronales convencionales, por otro lado en heridas quirúrgicas se hizo un mayor uso de datos para la predicción utilizando el machine learning. Contribuyendo a la orientación de futuras líneas de investigación.

En conclusión, los resultados de esta revisión de alcance evidencian el potencial de la IA en el manejo de heridas, especialmente el uso de datos clínicos y el análisis de imágenes para generar predicciones. Sin embargo, su implementación clínica requiere mayor rigor metodológico, evidenciando la necesidad de impulsar investigaciones prospectivas y multicéntricas con validación externa en entornos clínicos reales.

La Inteligencia Artificial representa una herramienta prometedora para fortalecer los sistemas de salud, reducir la carga laboral del personal y mejorar la eficacia terapéutica. No obstante, su aprovechamiento requiere la creación de guías y criterios regulatorios que garanticen su aplicación segura, ética y práctica.

11. Referencias

1. McNichol, Laurie L., editor., Ratliff, Catherine R., editor., Yates, Stephanie S., editor., Wound, Ostomy, and Continence Nurses Society. Core curriculum wound management. Wound, Ostomy, and Continence Nurses Society (WOCN) core curriculum. Second edition. Philadelphia: Wolters Kluwer; 2022. 40-46 p.
2. Conde E. Heridología. Primera edición. Estados Unidos: Desconocido; 2022. 18 p.
3. Grupo Nacional para el Estudio y Asesoramiento en Úlceras por Presión y Heridas Crónicas (GNEAUPP). Guía de práctica clínica Manejo y tratamiento de úlceras de extremidades inferiores - GNEAUPP [Internet]. 2021 [citado 2 de mayo de 2025]. Disponible en: [https://gneaupp.info/guia-de-practica-clinica-manejo-y-tratamiento-de-ulceras-de-extr emidades-inferiores/](https://gneaupp.info/guia-de-practica-clinica-manejo-y-tratamiento-de-ulceras-de-extr-emidades-inferiores/)
4. Lal BK. Venous ulcers of the lower extremity: Definition, epidemiology, and economic and social burdens. Semin Vasc Surg. marzo de 2015;28(1):3-5.
5. O'Donnell TF, Passman MA, Marston WA, Ennis WJ, Dalsing M, Kistner RL, et al. Management of venous leg ulcers: Clinical practice guidelines of the Society for Vascular Surgery® and the American Venous Forum. J Vasc Surg. 1 de agosto de 2014;60(2):3S-59S.
6. Ministerio de Sanidad, Servicios Sociales e Igualdad. Guía de Práctica Clínica sobre Úlceras en las Extremidades Inferiores [Internet]. [citado 5 de mayo de 2025]. Disponible en: https://www.sspa.juntadeandalucia.es/servicioandaluzdesalud/sites/default/files/sincfi les/wsas-media-pdf_publicacion/2020/Guia_de_cuidados_UA_def.pdf
7. Asociación Española de Enfermería Vascular y Heridas (AEEVH). Guía de Práctica Clínica en el Cuidado de las Úlceras Venosas [Internet]. [citado 5 de mayo de 2025]. Disponible en: <https://aeevh.org/wp-content/uploads/2020/04/Guia-de-Practica-Clinica-web.pdf>
8. International Working Group on the Diabetic Foot (IWGDF). Guía Práctica IWGDF 2023: Prevención y manejo del pie diabético [Internet]. 2023 [citado 5 de mayo de 2025]. Disponible en: <https://iwgdfguidelines.org/wp-content/uploads/2024/09/Guia-Espanol-IWGDF-2023.pdf>
9. Ma Luisa Paniagua Asensio. Prevención, clasificación y categorización de las lesiones relacionadas con la dependencia (LRD) [Internet]. 2020 [citado 5 de mayo de 2025]. Disponible en: <https://gneaupp.info/wp-content/uploads/2020/08/Paniagua-2020.-LRD.-Prevencio% CC%81n-clasificacio%CC%81n-y-categorizacio%CC%81n.pdf>

10. International Skin Tear Advisory Panel (ISTAP). La identificación y prevención de las heridas por desgarro cutáneo [Internet]. 2021 [citado 12 de mayo de 2025]. Disponible en: <https://iswcap.org/wp-content/uploads/2021/09/LA-IDENTIFICACION-Y-PREVENCIÓN.pdf>
11. Choi RY, Coyner AS, Kalpathy-Cramer J, Chiang MF, Campbell JP. Introduction to Machine Learning, Neural Networks, and Deep Learning. *Transl Vis Sci Technol*. 9(2):14.
12. IBM Corporation. ¿Qué es la Inteligencia Artificial (IA)? [Internet]. 2024 [citado 3 de abril de 2025]. Disponible en: <https://www.ibm.com/mx-es/think/topics/artificial-intelligence>
13. Chen M, Decary M. Artificial intelligence in healthcare: An essential guide for health leaders. *Healthc Manage Forum*. enero de 2020;33(1):10-8.
14. Massaron L. Deep learning. 1st edition. Hoboken, N.J: J. Wiley; 2019.
15. Piebpien P, Tansawet A, Pattanaprteep O, Pattanateepapon A, Wilasrusmee C, Mckay GJ, et al. Can machine learning models improve the prediction of surgical site infection in abdominal surgery than traditional statistical models? *J Int Med Res*. noviembre de 2024;52(11):3000605241293696.
16. McLean KA, Sgrò A, Brown LR, Buijs LF, Mountain KE, Shaw CA, et al. Multimodal machine learning to predict surgical site infection with healthcare workload impact assessment. *NPJ Digit Med*. 23 de febrero de 2025;8(1):121.
17. Sendilraj V, Pilcher W, Choi D, Bhasin A, Bhadada A, Bhadada SK, et al. DFUCare: deep learning platform for diabetic foot ulcer detection, analysis, and monitoring. *Front Endocrinol*. 2024;15:1386613.
18. Tusar MH, Fayyazbakhsh F, Zendehei N, Mochalin E, Melnychuk I, Gould L, et al. AI-Powered Image-Based Assessment of Pressure Injuries Using You Only Look once (YOLO) Version 8 Models. *Adv Wound Care*. 13 de marzo de 2025;
19. Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO). Inteligencia artificial: entre el mito y la realidad | El Correo de la UNESCO [Internet]. [citado 30 de abril de 2025]. Disponible en: <https://courier.unesco.org/es/articulos/inteligencia-artificial-entre-el-mito-y-la-realidad>
20. Yu X, Wu Z, Zhang N. Machine learning-driven discovery of novel therapeutic targets in diabetic foot ulcers. *Mol Med*. 14 de noviembre de 2024;30:215.
21. Rathore PS, Kumar A, Nandal A, Dhaka A, Sharma AK. A feature explainability-based deep learning technique for diabetic foot ulcer identification. *Sci Rep*. 25 de febrero de 2025;15(1):6758.

22. Lucas Y, Niri R, Treuillet S, Douzi H, Castaneda B. Wound Size Imaging: Ready for Smart Assessment and Monitoring. *Adv Wound Care*. noviembre de 2021;10(11):641-61.
23. Rippon MG, Fleming L, Chen T, Rogers AA, Ousey K. Artificial intelligence in wound care: diagnosis, assessment and treatment of hard-to-heal wounds: a narrative review. *J Wound Care*. 2 de abril de 2024;33(4):229-42.
24. Reifs Jiménez D, Casanova-Lozano L, Grau-Carrión S, Reig-Bolaño R. Artificial Intelligence Methods for Diagnostic and Decision-Making Assistance in Chronic Wounds: A Systematic Review. *J Med Syst*. 2025;49(1):29.
25. Pappachan JM, Cassidy B, Fernandez CJ, Chandrabalan V, Yap MH. The role of artificial intelligence technology in the care of diabetic foot ulcers: the past, the present, and the future. *World J Diabetes*. 15 de diciembre de 2022;13(12):1131-9.
26. van Boekel AM, van der Meijden SL, Arbous SM, Nelissen RGHH, Veldkamp KE, Nieswaag EB, et al. Systematic evaluation of machine learning models for postoperative surgical site infection prediction. *PLoS One*. 2024;19(12):e0312968.
27. Hamet P, Tremblay J. Artificial intelligence in medicine. *Metabolism*. abril de 2017;69S:S36-40.
28. Busaranuvong P, Agu E, Kumar D, Gautam S, Fard RS, Tulu B, et al. Guided Conditional Diffusion Classifier (ConDiff) for Enhanced Prediction of Infection in Diabetic Foot Ulcers. *IEEE Open J Eng Med Biol*. 2025;6:20-7.
29. Martinengo L, Olsson M, Bajpai R, Soljak M, Upton Z, Schmidtchen A, et al. Prevalence of chronic wounds in the general population: systematic review and meta-analysis of observational studies. *Ann Epidemiol*. 1 de enero de 2019;29:8-15.
30. Sen CK. Human Wound and Its Burden: Updated 2022 Compendium of Estimates. *Adv Wound Care*. diciembre de 2023;12(12):657-70.
31. Secretaría de Salud de Tabasco. Manual de procedimientos para el cuidado de heridas crónicas [Internet]. 2025 [citado 18 de mayo de 2025]. Disponible en: https://tabasco.gob.mx/sites/default/files/users/ssaludtabasco/869_0.pdf
32. Instituto Mexicano del Seguro Social (IMSS). Enfermería IMSS. Evaluación integral de heridas crónicas [Internet]. 2018 [citado 18 de mayo de 2025]. Disponible en: <https://www.medigraphic.com/pdfs/enfermeriaimss/eim-2018/eim182g.pdf>
33. Cohen JF, Bossuyt PMM. TRIPOD+AI: an updated reporting guideline for clinical prediction models. *BMJ*. 16 de abril de 2024;385:q824.
34. Skube SJ, Hu Z, Simon GJ, Wick EC, Arsoniadis EG, Ko CY, et al. Accelerating Surgical Site Infection Abstraction With a Semi-automated Machine-learning Approach. *Ann Surg*. 1 de julio de 2022;276(1):180-5.

35. Gutierrez-Naranjo JM, Moreira A, Valero-Moreno E, Bullock TS, Ogden LA, Zelle BA. -A machine learning model to predict surgical site infection after surgery of lower extremity fractures. *Int Orthop*. julio de 2024;48(7):1887-96.
36. Zhu Y, Simon GJ, Wick EC, Abe-Jones Y, Najafi N, Sheka A, et al. Applying Machine Learning Across Sites: External Validation of a Surgical Site Infection Detection Algorithm. *J Am Coll Surg*. junio de 2021;232(6):963-971.e1.
37. Chen W, Lu Z, You L, Zhou L, Xu J, Chen K. Artificial Intelligence-Based Multimodal Risk Assessment Model for Surgical Site Infection (AMRAMS): Development and Validation Study. *JMIR Med Inform*. 15 de junio de 2020;8(6):e18186.
38. Xiong C, Zhao R, Xu J, Liang H, Zhang C, Zhao Z, et al. Construct and Validate a Predictive Model for Surgical Site Infection after Posterior Lumbar Interbody Fusion Based on Machine Learning Algorithm. *Comput Math Methods Med*. 2022;2022:2697841.
39. Zhang Q, Chen G, Zhu Q, Liu Z, Li Y, Li R, et al. Construct validation of machine learning for accurately predicting the risk of postoperative surgical site infection following spine surgery. *J Hosp Infect*. abril de 2024;146:232-41.
40. Wu G, Cheligeer C, Southern DA, Martin EA, Xu Y, Leal J, et al. Development of machine learning models for the detection of surgical site infections following total hip and knee arthroplasty: a multicenter cohort study. *Antimicrob Resist Infect Control*. 2 de septiembre de 2023;12(1):88.
41. Lorenz WR, Holland AM, Sarac BA, Kerr SW, Wilson HH, Ayuso SA, et al. Development of Multicenter Deep Learning Models for Predicting Surgical Complexity and Surgical Site Infection in Abdominal Wall Reconstruction, a Pilot Study. *J Abdom Wall Surg JAWS*. 2025;4:14371.
42. Wang P, Liu L, Xie Z, Ren G, Hu Y, Shen M, et al. Explainable Machine Learning Models for Prediction of Surgical Site Infection After Posterior Lumbar Fusion Surgery Based on Shapley Additive Explanations. *World Neurosurg*. 1 de mayo de 2025;197:123942.
43. Chen KA, Joisa CU, Stem JM, Guillem JG, Gomez SM, Kapadia MR. Improved Prediction of Surgical-Site Infection After Colorectal Surgery Using Machine Learning. *Dis Colon Rectum*. 1 de marzo de 2023;66(3):458-66.
44. Lu K, Tu Y, Su S, Ding J, Hou X, Dong C, et al. Machine learning application for prediction of surgical site infection after posterior cervical surgery. *Int Wound J*. abril de 2024;21(4):e14607.
45. Craus-Miguel A, Fernández-Moreno A, Pablo-Leis AI, Romero-Hernández M, Munar M, Moyà-Alcover G, et al. Multicenter Study Protocol: Research on Evaluation

and Detection of Surgical Wound Complications with AI-Based Recognition (REDSCAR-Trial). *J Clin Med*. 24 de marzo de 2025;14(7):2210.

46. Zhuang Y, Dyas A, Meguid RA, Henderson WG, Bronsert M, Madsen H, et al. Preoperative Prediction of Postoperative Infections Using Machine Learning and Electronic Health Record Data. *Ann Surg*. 1 de abril de 2024;279(4):720-6.

47. Ohno Y, Mazaki J, Udo R, Tago T, Kasahara K, Enomoto M, et al. Preliminary Evaluation of a Novel Artificial Intelligence-based Prediction Model for Surgical Site Infection in Colon Cancer. *Cancer Diagn Progn*. 2022;2(6):691-6.

48. Chen T, Liu C, Zhang Z, Liang T, Zhu J, Zhou C, et al. Using Machine Learning to Predict Surgical Site Infection After Lumbar Spine Surgery. *Infect Drug Resist*. 2023;16:5197-207.

49. Yeo I, Klemt C, Robinson MG, Esposito JG, Uzosike AC, Kwon YM. The Use of Artificial Neural Networks for the Prediction of Surgical Site Infection Following TKA. *J Knee Surg*. mayo de 2023;36(6):637-43.

50. Fletcher RR, Schneider G, Bikorimana L, Rukundo G, Niyigena A, Miranda E, et al. The Use of Mobile Thermal Imaging and Deep Learning for Prediction of Surgical Site Infection. *Annu Int Conf IEEE Eng Med Biol Soc IEEE Eng Med Biol Soc Annu Int Conf*. noviembre de 2021;2021:5059-62.

51. Fletcher RR, Schneider G, Hedt-Gauthier B, Nkurunziza T, Alayande B, Riviello R, et al. Use of Convolutional Neural Nets and Transfer Learning for Prediction of Surgical Site Infection from Color Images. *Annu Int Conf IEEE Eng Med Biol Soc IEEE Eng Med Biol Soc Annu Int Conf*. noviembre de 2021;2021:5047-50.

52. Rhodes S, Sahmoud A, Jelovsek JE, Bretschneider CE, Gupta A, Hijaz AK, et al. Validation and Recalibration of a Model for Predicting Surgical-Site Infection After Pelvic Organ Prolapse Surgery. *Int Urogynecology J*. febrero de 2025;36(2):431-8.

53. Rochon M, Tanner J, Jurkiewicz J, Beckhelling J, Aondoakaa A, Wilson K, et al. Wound imaging software and digital platform to assist review of surgical wounds using patient smartphones: The development and evaluation of artificial intelligence (WISDOM AI study). *PloS One*. 2024;19(12):e0315384.

54. Shah RM, Khazanchi R, Bajaj A, Rana K, Malhotra S, Wolf JM. Using machine learning to identify risk factors for short-term complications following thumb carpometacarpal arthroplasty. *J Hand Microsurg*. diciembre de 2024;16(5):100156.

55. Curti N, Merli Y, Zengarini C, Starace M, Rapparini L, Marcelli E, et al. Automated Prediction of Photographic Wound Assessment Tool in Chronic Wound Images. *J Med Syst*. 2024;48(1).

56. Hüsters J, Hafer G, Heggemann J, Wiemeyer S, Przysucha M, Dissemond J, et al. Automatic Classification of Diabetic Foot Ulcer Images - A Transfer-Learning

Approach to Detect Wound Maceration. *Stud Health Technol Inform.* 14 de enero de 2022;289:301-4.

57. Chronic Ulcers Healing Prediction through Machine Learning Approaches: Preliminary Results on Diabetic Foot Ulcers Case Study. - EBSCO [Internet]. [citado 3 de noviembre de 2025]. Disponible en: <https://research-ebSCO-com.pbidi.unam.mx:2443/c/df24kt/viewer/html/xugbumjzbj>

58. Chan KS, Chan YM, Tan AHM, Liang S, Cho YT, Hong Q, et al. Clinical validation of an artificial intelligence-enabled wound imaging mobile application in diabetic foot ulcers. *Int Wound J.* enero de 2022;19(1):114-24.

59. Diabetic foot ulcer classification assessment employing an improved machine learning algorithm - EBSCO [Internet]. [citado 3 de noviembre de 2025]. Disponible en: <https://research-ebSCO-com.pbidi.unam.mx:2443/c/df24kt/viewer/html/ve63mqpctn>

60. Nanda R, Nath A, Patel S, Mohapatra E. Machine learning algorithm to evaluate risk factors of diabetic foot ulcers and its severity. *Med Biol Eng Comput.* agosto de 2022;60(8):2349-57.

61. Wang S, Xia C, Zheng Q, Wang A, Tan Q. Machine Learning Models for Predicting the Risk of Hard-to-Heal Diabetic Foot Ulcers in a Chinese Population. *Diabetes Metab Syndr Obes Targets Ther.* 2022;15:3347-59.

62. Mousa KM, Mousa FA, Mohamed HS, Elsayy MM. Prediction of Foot Ulcers Using Artificial Intelligence for Diabetic Patients at Cairo University Hospital, Egypt. *SAGE Open Nurs.* 2023;9:23779608231185873.

63. Xu J, Chen T, Fang X, Xia L, Pan X. Prediction model of pressure injury occurrence in diabetic patients during ICU hospitalization--XGBoost machine learning model can be interpreted based on SHAP. *Intensive Crit Care Nurs.* agosto de 2024;83:103715.

64. Zalluhoğlu C, Akdoğan D, Karakaya D, Güzel MS, Ülgü MM, Ardalı K, et al. Region-Based Semi-Two-Stream Convolutional Neural Networks for Pressure Ulcer Recognition. *J Digit Imaging.* 2024;37(2):801-13.

65. Nakagami G, Yokota S, Kitamura A, Takahashi T, Morita K, Noguchi H, et al. Supervised machine learning-based prediction for in-hospital pressure injury development using electronic health records: A retrospective observational cohort study in a university hospital in Japan. *Int J Nurs Stud.* julio de 2021;119:103932.

66. Popa AD, Gavril RS, Popa IV, Mihalache L, Gherasim A, Niță G, et al. Survival Prediction in Diabetic Foot Ulcers: A Machine Learning Approach. *J Clin Med.* 7 de septiembre de 2023;12(18):5816.

67. Seo S, Kang J, Eom IH, Song H, Park JH, Lee Y, et al. Visual classification of pressure injury stages for nurses: A deep learning model applying modern convolutional neural networks. *J Adv Nurs John Wiley Sons Inc.* 2023;79(8):3047-56.
68. Xiaoling W, Shengmei Z, BingQian W, Wen L, Shuyan G, Hanbei C, et al. Enhancing diabetic foot ulcer prediction with machine learning: A focus on Localized examinations. *Heliyon.* 15 de octubre de 2024;10(19):e37635.
69. Ge X, Du L, Zheng S, Shi A. Development and application of an intelligent pressure injury assessment system using AI image recognition. *Technol Health Care Off J Eur Soc Eng Med.* 9 de diciembre de 2024;9287329241296350.
70. Lau CH, Yu KHO, Yip TF, Luk LYF, Wai AKC, Sit TY, et al. An artificial intelligence-enabled smartphone app for real-time pressure injury assessment. *Front Med Technol.* 2022;4:905074.
71. Chen MY, Cao MQ, Xu TY. Progress in the application of artificial intelligence in skin wound assessment and prediction of healing time. *Am J Transl Res.* 2024;16(7):2765-76.
72. Tabja Bortesi JP, Ranisau J, Di S, McGillion M, Rosella L, Johnson A, et al. Machine Learning Approaches for the Image-Based Identification of Surgical Wound Infections: Scoping Review. *J Med Internet Res.* 18 de enero de 2024;26:e52880.

Anexo 1: Estrategia de búsqueda en bases de datos

Núm Búsqueda	Base de datos	Fecha	Términos y filtros	Resultados
Hx_Pubmed__1	PubMed	28/01/2025 29/06/2025	("Diabetic Foot"[Mesh]) OR "Pressure Ulcer"[Mesh]) AND "Artificial Intelligence"[Mesh]) OR "Machine Learning"[Mesh]) AND "Wound Infection"[Mesh] Sort by: Most Recent	54 59
Hx_Pubmed__2	PubMed	28/01/2025 29/06/2025	("Decision Support Systems, Clinical"[Mesh]) AND "Pressure Ulcer"[Mesh]) OR "Varicose Ulcer"[Mesh]) OR "Diabetic Foot"[Mesh]) AND "Predictive Value of Tests"[Mesh]) AND "Risk Assessment"[Mesh]	38 41
H_Pubmed_3	PubMed	28/01/2025	("Artificial Intelligence"[Mesh]) OR "Machine Learning"[Mesh] Filters: Systematic Review	2,619 61
Hx_Pubmed_4	PubMed	28/01/2025	("Chronic Wounds" OR "Diabetic Foot Ulcer" OR "Pressure Ulcer") AND ("Artificial Intelligence" OR "Machine Learning") AND ("Wound Infection")	5
Hx_Pubmed_5	PubMed	29/01/2025	Chronic wounds OR Diabetic foot ulcer OR Pressure ulcer AND artificial intelligence OR Machine Learning AND Wound Infection NOT surgical	54 64
Hx_Pubmed_6	PubMed	02/02/2025	("Artificial Intelligence"[Mesh])) OR "Generative Artificial Intelligence"[Mesh]) AND ("Wounds and Injuries/complications"[Mesh] OR "Wounds and Injuries/diagnosis"[Mesh] OR "Wounds and Injuries/diagnostic imaging"[Mesh] OR "Wounds and Injuries/prevention and control"[Mesh] OR "Wounds and Injuries/therapy"[Mesh])) AND "Wounds and Injuries"[Majr]) AND "Diabetic Foot/diagnosis"[Mesh]) AND "Diabetic Foot"[Majr]	1
7	Pubmed	29/06/2025	("artificial intelligence" OR "machine learning" OR "deep learning") AND ("wound" OR "wound healing" OR "surgical site infection" OR "diabetic foot ulcer" OR "chronic wounds") AND ("prediction" OR "risk assessment" OR "complications" OR "prognosis") Filters: in the last 10 years Sort by: Most Recent	406

Nombre de búsqueda	Base de datos	Fecha	Términos y filtros	Resultados
Hx_CINAHL_1	CINAHL	07/2025	<p>TI diabetic foot ulcer OR TI pressure ulcer AND TI artificial intelligence OR machine learning AND wound infection</p> <p>Limiters - Full Text; Publication Date: 20200101-20251231; English Language; Human; Publication Type: Meta Analysis, Randomized Controlled Trial, Systematic Review; Language: English, Spanish</p> <p>Expanders - Apply equivalent subjects</p> <p>Search modes - Find all my search terms</p>	165
Hx_CINAHL_2	CINAHL	07/2025	<p>TI (chronic wounds or chronic ulcers or non-healing wounds) OR TI diabetic foot ulcer OR TI pressure ulcer AND TI (artificial intelligence or ai or a.i.) OR TI machine learning AND TI wound infection</p> <p>Limiters - Full Text; Publication Date: 20200101-20251231; English Language; Human; Randomized Controlled Trials; Age Groups: All Adult; Language: English, Spanish</p> <p>Expanders - Apply equivalent subjects</p> <p>Search modes - Proximity</p>	44
Hx_CINAHL_3	CINAHL	07/2025	<p>(chronic wounds or chronic ulcers or non-healing wounds or venous leg ulcers or pressure wounds or diabetic foot ulcer) AND (artificial intelligence or ai or a.i. or machine learning or deep learning or artificial neural networks or algorithmic decision making)</p> <p>Limitadores - Texto completo; Fecha de publicación: 20200101-20251231; Artículo de investigación; Grupos de edad: All Adult; Idioma: English, Spanish</p> <p>Ampliadores - Aplicar palabras relacionadas</p> <p>Especificar por SubjectMajor: - artificial intelligence, generative wounds, chronic, wound measurement, wound care, wound assessment, venous ulcer, patient safety, motor skills, mortality, mobile applications, imaging, three-dimensional, hematoma, deep learning, clinical assessment tools, biological markers, amputation</p> <p>, treatment outcomes, intensive care units, image processing, computer assisted, boosting machine learning algorithms, algorithms, wounds and injuries, decision support systems, clinical, diabetic foot, prediction models, artificial intelligence, pressure ulcer, machine learning, risk assessment, adult: 19-44 years, aged: 65+ years. Modos de búsqueda - Buscar todos mis términos de búsqueda</p>	42

Hx_CINAHL_4	CINAHL	07/2025	(chronic wounds or chronic ulcers or non-healing wounds or venous leg ulcers or pressure wounds or diabetic foot ulcer) AND (artificial intelligence or ai or a.i. or machine learning or deep learning or artificial neural networks or algorithmic decision making) AND predictive Limitadores - Texto completo; Fecha de publicación: 20200101-20251231; Artículo de investigación; Humano; Tipo de publicación: Algorithm, Clinical Trial, Computer Program, Journal Article, Protocol, Randomized Controlled Trial; Grupos de edad: Adult: 19-44 years, Aged: 65+ years; Idioma: English, Spanish Ampliadores - Aplicar palabras relacionadas Modos de búsqueda - Buscar todos mis términos de búsqueda	20
Hx_CINAHL_5	CINAHL	07/2025	(chronic wounds or chronic ulcers or non-healing wounds or venous leg ulcers or pressure wounds or diabetic foot ulcer) AND (artificial intelligence or ai or a.i. or machine learning or deep learning or artificial neural networks or algorithmic decision making) AND (decision making or decision-making or decision making process or decision-making process) Limitadores - Texto completo; Fecha de publicación: 20200101-20251231; Artículo de investigación; Humano; Tipo de publicación: Algorithm, Clinical Trial, Computer Program, Journal Article, Protocol, Randomized Controlled Trial; Grupos de edad: Adult: 19-44 years, Aged: 65+ years; Idioma: English, Spanish Ampliadores - Aplicar palabras relacionadas Modos de búsqueda - Buscar todos los términos de búsqueda	7 no me salio
Hx_CINAHL_6	CINAHL	07/2025	XB (artificial intelligence or ai or a.i. or machine learning or deep learning) AND TI (surgical site infection or surgical wound infection or postoperative infection or ssi) OR TI dehiscd surgical wound Limiters - Full Text; Publication Date: 20200101-20251231; Human; Publication Type: Algorithm, Clinical Trial, Diagnostic Images, Journal Article, Protocol, Randomized Controlled Trial, Systematic Review; Age Groups: Adolescent: 13-18 years, Adult: 19-44 years, Middle Aged: 45-64 years; L ... Limiters - Full Text; Publication Date: 20200101-20251231; English Language; Research Article; Language: English, Spanish Search modes - Find all my search terms	9

Hx_CINAHL_7		07/2025	<p>(chronic wounds or chronic ulcers or non-healing wounds or venous leg ulcers or pressure wounds or diabetic foot ulcer) AND (artificial intelligence or ai or a.i. or machine learning or deep learning or artificial neural networks or algorithmic decision making)</p> <p>Limiters - Full Text; Publication Date: 20200101-20251231; English Language; Research Article</p> <p>Search modes - Find all my search terms</p>	63
-------------	--	---------	--	----

Anexo 2: Declaración de TRIPOD+AI

Section/Topic	Item	Development / evaluation ¹	Checklist item	Reported on page
TITLE				
<i>Title</i>	1	D,E	Identify the study as developing or evaluating the performance of a multivariable prediction model, the target population, and the outcome to be predicted	
ABSTRACT				
<i>Abstract</i>	2	D,E	See TRIPOD+AI for Abstracts checklist	
INTRODUCTION				
<i>Background</i>	3a	D,E	Explain the healthcare context (including whether diagnostic or prognostic) and rationale for developing or evaluating the prediction model, including references to existing models	
	3b	D,E	Describe the target population and the intended purpose of the prediction model in the context of the care pathway, including its intended users (e.g., healthcare professionals, patients, public)	
	3c	D,E	Describe any known health inequalities between sociodemographic groups	
<i>Objectives</i>	4	D,E	Specify the study objectives, including whether the study describes the development or validation of a prediction model (or both)	
METHODS				
<i>Data</i>	5a	D,E	Describe the sources of data separately for the development and evaluation datasets (e.g., randomised trial, cohort, routine care or registry data), the rationale for using these data, and representativeness of the data	
	5b	D,E	Specify the dates of the collected participant data, including start and end of participant accrual; and, if applicable, end of follow-up	
<i>Participants</i>	6a	D,E	Specify key elements of the study setting (e.g., primary care, secondary care, general population) including the number and location of centres	
	6b	D,E	Describe the eligibility criteria for study participants	
	6c	D,E	Give details of any treatments received, and how they were handled during model development or evaluation, if relevant	
<i>Data preparation</i>	7	D,E	Describe any data pre-processing and quality checking, including whether this was similar across relevant sociodemographic groups	
<i>Outcome</i>	8a	D,E	Clearly define the outcome that is being predicted and the time horizon, including how and when assessed, the rationale for choosing this outcome, and whether the method of outcome assessment is consistent across sociodemographic groups	
	8b	D,E	If outcome assessment requires subjective interpretation, describe the qualifications and demographic characteristics of the outcome assessors	
	8c	D,E	Report any actions to blind assessment of the outcome to be predicted	
<i>Predictors</i>	9a	D	Describe the choice of initial predictors (e.g., literature, previous models, all available predictors) and any pre-selection of predictors before model building	
	9b	D,E	Clearly define all predictors, including how and when they were measured (and any actions to blind assessment of predictors for the outcome and other predictors)	
	9c	D,E	If predictor measurement requires subjective interpretation, describe the qualifications and demographic characteristics of the predictor assessors	
<i>Sample size</i>	10	D,E	Explain how the study size was arrived at (separately for development and evaluation), and justify that the study size was sufficient to answer the research question. Include details of any sample size calculation	
<i>Missing data</i>	11	D,E	Describe how missing data were handled. Provide reasons for omitting any data	
<i>Analytical methods</i>	12a	D	Describe how the data were used (e.g., for development and evaluation of model performance) in the analysis, including whether the data were partitioned, considering any sample size requirements	
	12b	D	Depending on the type of model, describe how predictors were handled in the analyses (functional form, rescaling, transformation, or any standardisation)	
	12c	D	Specify the type of model, rationale ² , all model-building steps, including any hyperparameter tuning, and method for internal validation	
	12d	D,E	Describe if and how any heterogeneity in estimates of model parameter values and model performance was handled and quantified across clusters (e.g., hospitals, countries). See TRIPOD-Cluster for additional considerations ³	
	12e	D,E	Specify all measures and plots used (and their rationale) to evaluate model performance (e.g., discrimination, calibration, clinical utility) and, if relevant, to compare multiple models	
	12f	E	Describe any model updating (e.g., recalibration) arising from the model evaluation, either overall or for particular sociodemographic groups or settings	
	12g	E	For model evaluation, describe how the model predictions were calculated (e.g., formula, code, object, application programming interface)	
<i>Class imbalance</i>	13	D,E	If class imbalance methods were used, state why and how this was done, and any subsequent methods to recalibrate the model or the model predictions	
<i>Fairness</i>	14	D,E	Describe any approaches that were used to address model fairness and their rationale	
<i>Model output</i>	15	D	Specify the output of the prediction model (e.g., probabilities, classification). Provide details and rationale for any classification and how the thresholds were identified	

¹ D=items relevant only to the development of a prediction model; E=items relating solely to the evaluation of a prediction model; D,E=items applicable to both the development and evaluation of a prediction model

² Separately for all model building approaches.

³ TRIPOD-Cluster is a checklist of reporting recommendations for studies developing or validating models that explicitly account for clustering or explore heterogeneity in model performance (eg, at different hospitals or centres). Debnay et al, BMJ 2023; 380: e071018 [DOI: 10.1136/bmj-2022-071018]

<i>Training versus evaluation</i>	16	D,E	Identify any differences between the development and evaluation data in healthcare setting, eligibility criteria, outcome, and predictors
<i>Ethical approval</i>	17	D,E	Name the institutional research board or ethics committee that approved the study and describe the participant-informed consent or the ethics committee waiver of informed consent
OPEN SCIENCE			
<i>Funding</i>	18a	D,E	Give the source of funding and the role of the funders for the present study
<i>Conflicts of interest</i>	18b	D,E	Declare any conflicts of interest and financial disclosures for all authors
<i>Protocol</i>	18c	D,E	Indicate where the study protocol can be accessed or state that a protocol was not prepared
<i>Registration</i>	18d	D,E	Provide registration information for the study, including register name and registration number, or state that the study was not registered
<i>Data sharing</i>	18e	D,E	Provide details of the availability of the study data
<i>Code sharing</i>	18f	D,E	Provide details of the availability of the analytical code ¹
PATIENT & PUBLIC INVOLVEMENT			
<i>Patient & Public Involvement</i>	19	D,E	Provide details of any patient and public involvement during the design, conduct, reporting, interpretation, or dissemination of the study or state no involvement.
RESULTS			
<i>Participants</i>	20a	D,E	Describe the flow of participants through the study, including the number of participants with and without the outcome and, if applicable, a summary of the follow-up time. A diagram may be helpful.
	20b	D,E	Report the characteristics overall and, where applicable, for each data source or setting, including the key dates, key predictors (including demographics), treatments received, sample size, number of outcome events, follow-up time, and amount of missing data. A table may be helpful. Report any differences across key demographic groups.
	20c	E	For model evaluation, show a comparison with the development data of the distribution of important predictors (demographics, predictors, and outcome).
<i>Model development</i>	21	D,E	Specify the number of participants and outcome events in each analysis (e.g., for model development, hyperparameter tuning, model evaluation)
<i>Model specification</i>	22	D	Provide details of the full prediction model (e.g., formula, code, object, application programming interface) to allow predictions in new individuals and to enable third-party evaluation and implementation, including any restrictions to access or re-use (e.g., freely available, proprietary) ²
<i>Model performance</i>	23a	D,E	Report model performance estimates with confidence intervals, including for any key subgroups (e.g., sociodemographic). Consider plots to aid presentation.
	23b	D,E	If examined, report results of any heterogeneity in model performance across clusters. See TRIPOD Cluster for additional details ³ .
<i>Model updating</i>	24	E	Report the results from any model updating, including the updated model and subsequent performance
DISCUSSION			
<i>Interpretation</i>	25	D,E	Give an overall interpretation of the main results, including issues of fairness in the context of the objectives and previous studies
<i>Limitations</i>	26	D,E	Discuss any limitations of the study (such as a non-representative sample, sample size, overfitting, missing data) and their effects on any biases, statistical uncertainty, and generalizability
<i>Usability of the model in the context of current care</i>	27a	D	Describe how poor quality or unavailable input data (e.g., predictor values) should be assessed and handled when implementing the prediction model
	27b	D	Specify whether users will be required to interact in the handling of the input data or use of the model, and what level of expertise is required of users
	27c	D,E	Discuss any next steps for future research, with a specific view to applicability and generalizability of the model

Proveniente de: Collins GS, Moons KGM, Dhiman P, et al. BMJ 2024;385:e078378.
doi:10.1136/bmj-2023-078378

Anexo 3: Tablas de resumen heridas crónicas

#	Título	País	Año	Población y tipo de herida	Diseño de investigación	Metodología	Programa de IA que utilizaron	Propósito de la IA	Resultados	Limitaciones	Conclusiones
1	Automated Prediction of Photographic Wound Assessment Tool in Chronic Wound Images	Italia	2024	474 pacientes geriátricos 612 imágenes de heridas crónicas en distintas fases de cicatrización	Observacional Retrospectivo con validación interna	Captura de imágenes con celular donde se hizo segmentación automática de con red neuronal (CNN) Se extrajeron características (color, textura y forma) para un modelo de regresión Lasso para predecir el puntaje PWAT	Red neuronal convencional (CNN) Regresión Lasso Colores HSV	Calculo automatico de PWAT (photographic wound assesment tool) para no hacer una evaluación manual	El sistema es útil Correlación Spearman: 0.85 entre el modelo automático y la evaluación clínica	Solo se validó en un centro Depende de la cámara y la iluminación No se ha probado en otras poblaciones y no había casos muy severos	IA: evalúa de forma objetiva las heridas crónicas y puede funcionar de apoyo Este modelo ya es utilizado en ese hospital.
2	Automatic Classification of Diabetic Foot Ulcer Images - A Transfer-Learning Approach to Detect Wound Maceration	Alemania	2022	434 imágenes de úlceras de pie diabético (DFU), clasificadas por especialistas según presencia o ausencia de maceración. No pacientes directamente, sino imágenes clínicas.	Experimental retrospectivo Validación interna	Imágenes de úlceras fueron recolectadas en un centro especializado. Se anotaron manualmente con cajas delimitadoras. Se clasificaron según la maceración. Se dividieron en 75% entrenamiento y 25% validación. Se usó transferencia de aprendizaje con MobileNetV1 y técnicas de data augmentation.	Deep Learning: red neuronal convolucional (CNN) MobileNetV1 con transferencia de aprendizaje y fine-tuning en TensorFlow.	Clasificar automáticamente e imágenes de DFU según presencia de maceración para apoyar documentación clínica y reducir el tiempo de registro médico.	F1-score: 0.71- Recall (sensibilidad): 0.69 Precisión (PPV): 0.67 Exactitud: 0.69- AUC: 0.78	Solo imágenes de un centro Sin validación externa Modelo no probado con imágenes que tengan cremas o apósitos Precisión insuficiente para implementación clínica automática	El modelo puede ser útil como sistema de apoyo a la documentación, pero necesita mejoras para uso clínico automático. Puede rellenar datos y permitir que el médico confirme o rechace.

3	Chronic Ulcers Healing Prediction through Machine Learning Approaches: Preliminary Results on Diabetic Foot Ulcers Case Study	Italia Francia	2025	1766 registros de pacientes con úlceras crónicas específicamente en pie diabético	Retrospectivo observacional con validación cruzada	Usaron un dispositivo Wound Viewer para recolectar imágenes y datos clínicos durante >3 visitas por paciente Clasificando el estado de la herida entre empeora y mejora	Machine Learning (KNN, SVM, RF, Naïve Bayes, XGBoost, LightGBM) y Deep Learning (LSTM, FCNN).	Predecir si la herida del pie diabético va a mejorar o empeorar en la siguiente consulta, para apoyar decisiones clínicas tempranas	El modelo LSTM de 3 capas obtuvo: AUC = 0.85, Precisión = 79%, Recall = 80%, F1-score = 80%. Las variables más influyentes fueron área, profundidad y tejido granulado	Dataset no balanceado, sin validación externa, especificidad menor, cómputo no optimizado aún, no está implementado en tiempo real clínico	La IA puede predecir la evolución de úlceras crónicas con alta precisión. Tiene potencial para usarse en la práctica médica con validación futura.
4	Clinical validation of an artificial intelligence-enabled wound imaging mobile application in diabetic foot ulcers	Singapur	2022	28 pacientes con úlceras del pie diabético (547 imágenes analizadas)	Prospectivo transversal	Comparación entre medición manual por enfermeras (papel milimetrado) y medición digital con la app C4W (en iPhone 8+, 11 Pro, XS). Se evaluó confiabilidad intra- e interevaluador usando ICC.	App C4W con IA embebida de visión computacional (clasifica bordes y tejidos en imágenes clínicas, ya entrenada previamente)	Medir automáticamente e las dimensiones de las heridas en pacientes con DFU, para seguimiento clínico objetivo y rápido	Intra-rater ICC: 0.933–0.994. Inter-rater ICC vs enfermera: 0.825–0.934. App sobreestimó ligeramente el área vs método manual.	No evaluó profundidad ni características cualitativas (exudado, esfacelo), ni validación en otros dispositivos más antiguos.	Alta confiabilidad en mediciones. Herramienta útil como complemento para seguimiento de heridas DFU. Aún requiere validación más amplia.
5	DFUCare: deep learning platform for diabetic foot ulcer detection, analysis, and monitoring	Estados Unidos, India	2024	Pacientes con úlceras de pie diabético	Estudio de desarrollo y validación del modelo de IA	Se usó algoritmos de visión por computadora y aprendizaje profundo para la detección, clasificación y análisis de heridas	DFUCare plataforma con Deep Learning	Localización, clasificación y análisis de úlceras de pie diabético	F1-score de 0.80 y mAP de 0.861 para localización de heridas; 79.76% de precisión en clasificación de infección y 94.81% en isquemia	Falta de diversidad en la base de datos, necesidad de validación con algunas imágenes en pacientes de distintos grupos étnicos y condiciones de iluminación	La plataforma presentada ha demostrado buen monitoreo de las heridas y podría ayudar al monitoreo de pacientes y que estos tengan mejores resultados

6	Diabetic foot ulcer classification assessment employing an improved machine learning algorithm	India	2025	Imágenes de úlceras de pie diabético extraídas de bases de datos públicas (DFUC2020 dataset) y privadas recolectadas localmente. Clasificación automática de tipo de lesión.	Experimental Validación interna	Preprocesamiento de imágenes (resizing, normalización), extracción de características, clasificación por Random Forest mejorado (IM-RF). Dividido en 70% entrenamiento y 30% prueba. Validación con 10-fold cross validation.	Machine Learning tradicional: Improved Random Forest (IM-RF), comparado con Support Vector Machine, K-Nearest Neighbors y Artificial Neural Networks.	Clasificar automáticamente e imágenes de DFU en categorías (infectadas, no infectadas, maceradas, etc.), con mayor precisión y robustez que modelos convencionales.	IM-RF alcanzó: precisión = 94.1%, sensibilidad = 93.6%, especificidad = 95.3%, AUC = 0.948. Superó a los demás modelos comparados.	No probado en entorno clínico. No se especifica si el dataset incluye suficiente variabilidad étnica o ambiental. Validación solo interna.	El modelo IM-RF es una mejora significativa en clasificación de imágenes de DFU. Puede usarse para apoyo diagnóstico automatizado y reducir carga clínica.
7	Guided Conditional Diffusion Classifier (ConDiff) for Enhanced Prediction of Infection in Diabetic Foot Ulcers	Estados Unidos	2024	Pacientes con úlceras de pie diabético	Estudio observacional con validación cruzada de 5 pliegues	Recopilación de información con imágenes de UPD recopiladas del hospital Universitario de Lancashire Uso de modelo de difusión guiada ConDiff que aplica ruido gaussiano con un proceso de descomposición de imágenes sintéticas	Aprendizaje profundo CNN	Detección y predicción de infecciones en úlceras de pie diabético a partir de imágenes tomadas en el punto de atención	Modelo de difusión de eliminación de ruido en fotos de UPD Comparación vs SOTA Utilidad para limitar la brecha de la distancia Sensibilidad de 85.4% y especificidad 74.4%	Alto costo computacional No predice la progresión futura de la infección, solo el estado actual	ConDiff mejora la detección de infecciones en DFU, superando los modelos existentes y mostrando mejor generalización en la clínica
8	Machine learning algorithm to evaluate risk factors of diabetic foot ulcers and its severity	India	2022	Px con y sin pie diabético, de entre 18 y 80 años 160 px → 80 de ellos con DM2 y úlceras de pie diabético (UPD) y 80 con DM2 sin UPD	Observacional, prospectivo y unicéntrico	2 etapas: 1era, se dividió a px con DM2 sin UPD y pacientes con DM2 con UPD. 2da, se entrenaron modelos de aprendizaje automático para clasificar los tipos de úlceras en el grupo con UPD	Aprendizaje automático: RF, SVM-Poly K, KNN, NB, Stacking C, AdaBoost	Construir modelos de predicción para la clasificación del tipo de úlcera	SVM-PolyK MCC (0,875) y F-measure (0,938) superior a otros algoritmos. Al combinar los resultados de SVM-PolyK y RF (los 2 algoritmos con mejor rendimiento) mediante Stacking C: AUC (MCC: 0,875, F-measure: 0,938 y AUC: 0,970).	Muestra pequeña No multicéntrica	Se realizó un sistema discriminatorio con el uso de algoritmos de aprendizaje automático identificando FR que ayudaban a la predicción de UPD, también podría discriminar entre los grados de úlceras según su gravedad.

									RF alcanzó una sensibilidad del 95 %, SVM-PolyK especificidad del 93,8% AdaBoost = 86,3%, Bagging = 88,1%		
9	Machine Learning Models for Predicting the Risk of Hard-to-Heal Diabetic Foot Ulcers in a Chinese Population	India	2022	160 pacientes (80 con DFU y 80 sin DFU).	Observacional Prospectivo	Se recolectaron datos clínicos y bioquímicos (29 características, incluyendo IL-10, ApoA1, HbA1C). Comparación de diferentes algoritmos	Naive Bayes, AdaBoost, Bagging, RF, SVM, KNN	Identificar factores de riesgo y clasificar el grado de severidad de la úlcera (clasificación Wagner).	Mejor rendimiento con SVM-PolyK + Random Forest en Stacking (AUC 0.97, sensibilidad 95%).	Base de datos pequeña y necesidad de validación externa.	El uso de múltiples algoritmos en conjunto mejora la predicción y clasificación de DFU
10	Prediction of Foot Ulcers Using Artificial Intelligence for Diabetic Patients at Cairo University Hospital, Egypt.	Egipto	2023	200 pacientes con DM tipo 2 (82 con úlceras en pie y 118 sin úlceras). 75% entrenamiento y 25% prueba.	Casos y Controles	Datos demográficos, clínicos e imágenes del pie. Se realizaron 2 algoritmos de IA (red neuronal artificial ANN y árbol de decisión DT) implementados con MATLAB. árbol de decisión utilizando características de imágenes (13 variables) y datos clínicos (6 variables seleccionadas). SPSS ver 16	ANN, DT MATLAB.	Predicción temprana de riesgo de úlcera diabética.	ANN precisión de 0,97 en comparación con 0,93 con el DT La sensibilidad de la ANN fue mejor que la del DT; la especificidad y el VPP fueron de 1 y 0,99, respectivamente, superiores a los del DT. La ANN supera al DT en cuanto al VPN. Finalmente, se observa un aumento de 0,05 en la puntuación F con la ANN en comparación con el DT	Tamaño de muestra reducido y sólo datos de un único hospital.	El ANN es una herramienta muy prometedora para identificar pacientes con riesgo de úlcera.
11	Prediction model of pressure injury occurrence in diabetic patients	China	2024	503 pacientes con DM hospitalizados en UCI Tipo de herida:	Cohorte Retrospectivo	Recolección de variables clínicas y selección de las variables: LASSO 75% entrenamiento y	XG Boost	Predecir la probabilidad temprana de desarrollar lesiones por	AUC: Entrenamiento: 0.896 Validación: 0.835 Especificidad:	No multicéntrico Tamaño de muestra limitado Sin seguimiento a largo plazo	XGBoost + SHAP proporciona un modelo predictivo robusto y explicable para

	during ICU hospitalization — XGBoost machine learning model can be interpreted based on SHAP			LPP durante estancia en UCI Incidencia: 33.8% (170 casos con PI)		25% validación Software: R (paquetes glmnet, caret, xgboost, shapviz)		presión en pacientes con diabetes en UCI	91.6% Sensibilidad: 52.4% F1-score: 0.62 en validación SHAP muestra impacto individual de cada variable (mayor tiempo en UCI y ventilación aumentan riesgo). Aunque SHAP mejora interpretabilidad, sigue siendo complejo para clínicos sin formación en IA		riesgo de LPP en DM en UCI. Permite decisiones proactivas y priorización de cuidados preventivos. Se recomienda validación externa y adaptación en sistemas EHR.
12	Region-Based Semi-Two-Stream Convolutional Neural Networks for Pressure Ulcer Recognition	Turquía	2024	Px con úlceras por presión, imágenes obtenidas del hospital Pursaklar State Hospital (Ankara, Turquía), principalmente en px en UCI 1202: 907 entrenamiento y 295 de prueba.	Retrospectivo experimental	Clasificación según EPUAP (estadio 1 a 4). Se desarrollaron y probaron 12 arquitecturas de CNN. Semi-Two-Stream CNN que combina imagen completa más una región de interés (ROI). Se realizó transfer learning sobre modelos pre entrenados en ImageNet. Entrenamiento: 200 épocas, batch size 64, SGD optimizer, tasa aprendizaje inicial 0.01, reducción progresiva.	CNN Transfer Learning, ROI Pooling	Clasificar automáticamente e las úlceras por presión en 4 estadios según EPUAP, para facilitar el diagnóstico temprano y reducir costos y complicaciones.	Benchmark con modelos estándar: Precisión promedio 77-79% (VGG16 y DenseNet mejores). Modelo propuesto Semi-Two-Stream CNN: Exactitud: 93% Precisión: 93% Recall: 92% F1-score: 93%	Bde datos monocéntrico (solo un hospital). Población homogénea Imágenes tomadas con cámaras simples, sin control de luz ni calibración. Dataset pequeño en comparación con otros problemas de IA	Nuevo dataset PU (con imágenes completas y ROIs) contribuye al campo. La propuesta Semi-Two-Stream mejora los modelos estándar. La IA puede ser útil como herramienta de soporte para el diagnóstico de úlceras por presión. Requiere de validación externa
13	Supervised machine learning-based	Japón	2021	Uso de 2.940 registros relacionados con el	Cohorte observacional retrospectivo	Se utilizan 4 clasificadores principales para	Aprendizaje automático	Predecir lesiones por presión durante	395 px desarrollaron úlceras por presión	Al abarcar todos los ingresos hospitalarios, una	Utilizando datos de la Historia Clínica

	prediction for in-hospital pressure injury development using electronic health records: A retrospective observational cohort study in a university hospital in Japan			manejo de las lesiones por presión de 2.192 pacientes que desarrollaron lesiones por presión recientemente Total de 75.353 px identificados		probar: regresión logística, bosque aleatorio (RF), máquina de vectores de soporte lineal (Linear SVM) y potenciación de gradiente extrema (XGBoost). Uso de AUC para evaluar el rendimiento predictivo. Datos originales divididos, 70 % de datos de entrenamiento y un 30 % de datos de prueba		la hospitalización	durante la hospitalización. La proporción de úlceras por presión adquiridas en el hospital fue del 0,52 %. El Modelo XGBoost logró la mayor sensibilidad (0,78±0,03) y AUC (0,80± 0,02) entre los cuatro tipos de clasificadores. Las curvas ROC revelaron el mejor rendimiento predictivo en el modelo XGBoost utilizando los datos de prueba, con un AUC de 0,80	incidencia del 0,5 % resultó en un valor predictivo positivo significativamente bajo, de alrededor del 1,5 % La validez externa del modelo propuesto es baja, ya que se ajusta al hospital donde se recopilaron los datos.	Electrónica recopilados por el personal de enfermería, hemos creado una técnica de aprendizaje automático para construir modelos de predicción sobre el desarrollo de lesiones por presión durante la hospitalización en un hospital universitario
14	Survival Prediction in Diabetic Foot Ulcers: A Machine Learning Approach	Rumania	2023	635 px diagnosticadas con DM tipo 1 y tipo 2 con UPD asociadas monitoreados hasta su fallecimiento o última fecha documentada	Retrospectivo observacional	Base de datos en colaboración con el servicio de estadística del hospital. Los datos recopilados incluyeron variables continuas, categóricas. El resultados de los modelos diseñados deben predecir variables binarias estimando si el px morirá en los 5 o 10 años sig a la hospitalización	Método de Aprendizaje Automático (ML) Percepción multicapa (MLP)	Predecir la mortalidad a largo plazo en pacientes diabéticos con úlceras del pie diabético	1. Modelo de mortalidad a 5 años Conjunto de entrenamiento → AUC 0,755, VPN 0.6442 y PPV 0.86 Conjunto de prueba → AUC 0,73, VPN 0.5962 y PPV 0.8133 2. Modelo de mortalidad a 10 años Conjunto de entrenamiento → AUC 0,8095, VPN 0.7170 y PPV 0.7905 Conjunto de prueba → AUC 0.7339, VPN 0.6111 y PPV 0.7808	Falta de inclusión del enfoque terapéutico específico, ya sea médico o quirúrgico, utilizado para el tratamiento de las úlceras del pie diabético (UPD) Disponibilidad de historias clínicas electrónicas que no incluían la causa de muerte	Es posible predecir con buena precisión el riesgo de muerte en pacientes con UPD en los siguientes 5 o 10 años mediante predictores clínicos y biológicos no invasivos y de bajo costo
15	Visual	Korea	2023	2464 imagenes de	Retrospectivo	Un cirujano plástico y	CNN	Clasificación de	EfficientNet, al	La clasificación	Se creó un

	classification of pressure injury stages for nurses: A deep learning model applying modern convolutional neural networks			LPP 80% entrenamiento, 10% validación y 10% prueba	Observacional	dos enfermeras especialistas etiquetaron un conjunto de imágenes de LPP (Estadio 2,3,4, profundas, no estadificar y DAI) Uso de Python para construir el modelo de DL. Aprendizaje por transferencia	Comparado con VGG16, ResNet50, ResNet 152, DenseNet 201, and EfficientNet. VGG16	las etapas de las lesiones por presión en fotografías del mundo real y comparar su desempeño con el de las enfermeras clínicas	clasificar imágenes de lesiones por presión, se calculó el puntaje macro F1. Se informó que el rendimiento promedio de dos enfermeras experimentadas fue de 0,8941. como 0,8781.	de los estadios podía no ser exacta debido a que solo se realizó con imágenes No se consideró estadio 1 ni en mucosas	modelo que estadifica sin embargo se necesitan hacer estudios prospectivos para evaluar su funcionalidad
16	A feature explainability-based deep learning technique for diabetic foot ulcer identification	India	2025	Personas con diabetes y riesgo de tener úlceras neuropáticas N: 1050 parches de piel 540 = piel normal 510 = úlceras	Desarrollo y validación de modelo de IA Estudio observacional de cohorte retrospectivo	Evaluación de seis modelos avanzados de aprendizaje profundo utilizando técnicas de explicabilidad como SHAP, LIME y Grad-CAM	Utilizaron deep learning (DL): Modelo que ha sido entrenado para extraer datos visuales de pies diabéticos CNN: algoritmo que utiliza mapas de color o segmentación de regiones para pronostico de heridas DFU XAI:incorpora tres enfoques para reducir sesgos, aumento de datos, medidas de rendimiento	Detección, localización y clasificación de úlceras de pie diabético	SNN alcanzó 98.76% de precisión, 99.3% de precisión (precision), 97.7% de recall, 98.5% de F1-score y 98.6% AUC	Funciona bien para úlceras neuropáticas pero no para otras condiciones de la piel Riesgos de confidencialidad Seguridad del modelo por privacidad de datos	El modelo SNN con técnicas de explicabilidad mejora la confianza en la IA en entornos clínicos para la detección de úlceras de pie diabético

17	Enhancing diabetic foot ulcer prediction with machine learning: A focus on Localized examinations	China	2024	Pacientes con diabetes mellitus y riesgo de úlceras de pie diabético	Estudio observacional	Evaluar el riesgo de presentar pie diabético Base de datos: 566 casos Grupo de entrenamiento 453 Grupo de validación 113	MGWC SVM RF aprendizaje automático automatizado	Predecir el riesgo de desarrollar úlceras de pie diabético a partir de factores clínicos	AUC: 88.48% (74.44%–97.83%) Sensibilidad: 87.23% (63.33%–100.00%) Especificidad: 87.43% (70.00%–100.00%) Precisión: 87.33% (76.66%–95.00%) Valor predictivo positivo: 87.69% (75.00%–100.00%) Valor predictivo negativo: 87.70% (71.79%–100.00%)	Muestra relativamente pequeña Modelo basado en Auto ML	El modelo de predicción basado en AutoML mostró alta precisión y potencial clínico para la detección
18	Machine learning-driven discovery of novel therapeutic targets in diabetic foot ulcers	China	2024	15 px con infección en úlcera de pie diabético a la semana 0 post r tratamiento (control) a las 8 semanas post infección y con tto (grupo experimental)	Estudio observacional con análisis bioinformático	Uso de R Base de datos GEO Prueba LASSO SVM-RFE, ROC,	Aprendizaje automático	Identificar nuevos objetivos terapéuticos y biomarcadores para el diagnóstico y tratamiento	Identificación se SCUBE1 y RNF103-CHMP3 como posibles biomarcadores de la UPD	Se requiere mayor evidencia	SCUBE1 y RNF103-CHMP3 pueden ser usados a futuro como biomarcadores que sean usados como objetivo terapéutico
19	Development and application of an intelligent pressure injury assessment system using AI image recognition	China	2024	108 paciente mayores de 18 años con estancia de más de 48 hrs en UCI con lesiones por presión	Desarrollo y validación de modelo de IA experimental	Grupo experimental de 54 pacientes, con un 20 % adicional para compensar la deserción. Grupo de control: evaluación estándar y grupo experimental: sistema inteligente	CNN	Medir el tamaño de la herida, la categorización del tejido y la intensidad de la úlcera entre estadios I a IV. Integración de los datos a los registros clínicos de la institución	El sistema de IA alcanzó una tasa de precisión del 88%, una tasa de recuperación del 92% y una puntuación F1 del 90%. El grupo experimental demostró una mayor precisión en la evaluación (90	Evaluar el rendimiento del sistema en diversos entornos clínicos y con diferentes tipos de heridas. Se requiere integrar a otras tecnologías emergentes y adaptación para la telemedicina	Los sistemas de inteligencia artificial con tecnología de reconocimiento de imágenes mejoran la precisión y la eficacia de la evaluación de la heridas.

									% en comparación con el grupo control (81,2 %) Reducción de tiempo de evaluación en uso de IA		
20	AI-Powered Image-Based Assessment of Pressure Injuries Using You Only Look Once Version 8 (YOLOv8) Models	Estados Unidos	2025	Uso de 720 imágenes que fueron estratificadas según el NPIAP. Conjunto de imágenes independientes	Desarrollo y validación de modelo de IA Observacional	Se probó el modelo en un conjunto de imágenes independiente que no se utilizó en el entrenamiento. Las imágenes fueron evaluadas de forma independiente por médicos	IA YOLOv8 utiliza la red troncal CSPDarknet53	Detección y estadificación de las LPP utilizando YOLOv8.	Se utilizaron 5 diferentes modelos YOLOv 8 Precisión Etapa 2:0,64. DTI: 0,87 No clasificables:0,92 YOLOv8x logró altas precisiones para DTI (0,84), lesiones no clasificables (0,91) y Etapa 1 (0,86), pero tuvo problemas con la Etapa 2 (0,64).	Los modelos más grandes, como YOLOv8l y YOLOv8x, requieren conjuntos de datos más grandes para aprovechar al máximo su capacidad y mejorar el rendimiento.	YOLOv8 y los modelos de conjunto permiten mejorar la precisión y la robustez de la estadificación de las lesiones por presión, ofreciendo una herramienta fiable para la toma de decisiones clínicas.
21	An artificial intelligence-enabled smartphone app for real-time pressure injury assessment	Hong Kong	2022	Lesiones por presión en adultos mayores. Se usaron imágenes públicas y de casos clínicos reales	Estudio piloto Validación técnica de app (se simularon escenarios y también hubo reales como parte del testeo)	Desarrollo de app Android con YOLOv4 Entrenado con imagenes publicas n: 190 Validación con 144 fotos y con 10 fotos impresas simuladas en un brazo	YOLOv4 (you only look once version 4): redes neuronales y deep learning Se implementaron con Darknet y Tensor Flow LITE	Detectar en tiempo real las lesiones por presión y clasificarlas	Precisión de 63.2% Especificidad de 85.1 y 100% Sensibilidad de 37-76.7% Prueba de precisión 80-90%	Se usaron pocas imagenes y tambien fotos impresas No se usaron pacientes reales Solo era version Android	La app puede promover el diagnóstico temprano y remoto de las lesión por presión Requiere una validación clínica y ampliación de plataforma

Anexo 4: Tablas de resumen heridas quirúrgicas

#	Título	País	Año	Población y tipo de herida	Diseño de investigación	Metodología	Programa de IA que utilizaron	Propósito de la IA	Resultados	Limitaciones	Conclusiones
1	Accelerating Surgical Site Infection Abstraction With a Semi-automated Machine-learning Approach	Estados Unidos (UCLA, UCSF, American College of Surgeons)	2022	Pacientes sometidos a cirugía (n:11,320) con posibles infecciones de sitio quirúrgico evaluadas en los años 2011 a 2015 y registradas en NSQUIP	Cohorte Retrospectivo	Regresión con modelo Lasso Se usaron datos estructurados y notas clínicas con NLP Se definieron puntos de corte para reducir la revisión manual	Regresión logística penalizada con Lasso + procesamiento de lenguaje natural (NLP-PIER)	Detectar infecciones de sitio quirúrgico con alta especificidad para reducir revisión manual	Reducción del >90% Sensibilidad entre 82% y 89.4% Falsos negativos <1% Especificidad y valor predictivo del 100%	Necesidad de calibración con el tiempo Depende del acceso completo a datos clínicos	La IA semi automatizada ayuda a reducir castor y carga de trabajo en programas como NSQIP con buena capacidad de detección de SSI
2	A machine learning model to predict surgical site infection after surgery of lower extremity fractures	Estados Unidos (San Antonio Texas)	2024	1,579 pacientes con fracturas de extremidad inferior (tobillo, pierna) Heridas quirúrgicas postoperatorias	Retrospectivo Cohorte Validación interna	Registros institucionales para analizar 14 variables predictoras Selección de predictores con MARS Se entrenaron 4 modelos de ML	Redes neuronales Modelo Lineal generalizado Naive Bayes Selección de variables con MARS	Predicción de sitio infecciones después de cirugía	Modelo con mejor desempeño (lineal generalizado potenciado) AUC 77.4% Sensibilidad 83.3% Especificidad 62.4% Brier Score 5.1%	Estudio retrospectivo, sin validación externa Posibilidad de sesgos Aplicación limitada en diferentes poblaciones	El modelo puede estratificar el riesgo de infección quirúrgica en cirugía ortopédica pero requiere validación externa

3	Applying Machine Learning Across Sites: External Validation of a Surgical Site Infection Detection Algorithm	Estados Unidos (Universidad de Minnesota, Universidad de California y San Francisco)	2021	Pacientes quirúrgicos con riesgo o infección de sitio quirúrgico 8,883 Minnesota 1,473 San Francisco	Cohorte Retrospectivo Validación externa	Extracción de datos del expediente electrónico Se usó regresión lógica personalizada (Lasso) Se seleccionaron las variables mediante algoritmo de PC-Simple, Validación cruzada y bootstrapping	Regresión logística penalizada con Lasso + filtrado de variables con algoritmo PC simple	Detectar de forma automática infecciones postoperatorias para reducir el trabajo manual	AUC en validación externa con .804 (superficial), 0.905 (órgano/espacio) Reducción significativa de la revisión manual	Solo se realizó en dos hospitales por el mismo sistema de expediente electrónico Existe el riesgo de deterioro del modelo con el tiempo Falta de validación en hospitales comunitarios	El módulo se generaliza bien y reduce la carga de revisión pero es necesario adaptarlo a otros hospitales
4	Artificial Intelligence-Based Multimodal Risk Assessment Model for Surgical Site Infection (AMRAMS): Development and Validation Study	China	2020	21,611 pacientes que estuvieron en cirugía general, ginecología, urología y ortopédica	Cohorte Estudio retrospectivo Validación interna y externa	Se recuperaron datos de registros clínicos en un hospital en China, usando datos previos al 2018 para la validación interna y el resto para validación externa. incluyendo datos geográficos, resultados de laboratorio preoperatorios y notas preoperatorias de texto libre	LASSO Random Forest GBDT CNN Técnica de fast text	Predecir el riesgo de infecciones quirúrgicas preoperatorias utilizando datos clínicos	AUROC hasta 0.889 en validación interna (CNN) Validación externa: 0.879 (self attention)	Datos de un solo hospital Falta de datos por lo que puede que no se hayan detectado todas las SSI Recursos limitados para modelos más complejos	La IA mejora la predicción de SSI cuando se utilizan modelos multimodales

5	Can machine learning models improve the prediction of surgical site infection in abdominal surgery than traditional statistical models?	Tailandia	2024	12,596 cirugías abdominales (gastrointestinal, colorrectal, hernias). Tipo: heridas quirúrgicas, algunas contaminadas o limpias con cuerpo extraño.	Cohorte retrospectivo Validación externa	Revisión externa del modelo SENIC (estadístico clásico) Comparación con 4 modelos de machine learning usando datos hospitalarios	Machine Learning clásico Naive, bayes DT RF XG Boost	Predecir riesgo de infección de sitio quirúrgico (SSI)	El modelo clásico actualizado (SENIC-LR) tuvo mejor desempeño: C-statistic = 0.768. - ML más preciso: Naïve Bayes con C-statistic = 0.679. Todos los modelos ML tuvieron problemas de overfitting. - Factores más importantes: tiempo de cirugía >2h, transfusión sanguínea, cirugía abierta, diabetes, procedimientos múltiples.	Solo se aplicó a tipos específicos de cirugía abdominal Pocos casos de infección (3.21%) vs. muchas variables = riesgo de sobreajuste No se hicieron ajustes por clase desbalanceada	El modelo actualizado SENIC y Naive Bayes son útiles para predecir SSI pero aún requiere evaluación prospectiva en más hospitales antes de aplicarse clínicamente.
6	Construct and Validate a Predictive Model for Surgical Site Infection after Posterior Lumbar Interbody Fusion Based on Machine Learning Algorithm	China	2022	584 pacientes con enfermedad lumbar degenerativa sometidos a cirugía de fusión intersomática lumbar posterior. De ellos, 33 desarrollaron infección del sitio quirúrgico (SSI).	Cohorte Estudio retrospectivo observacional	Recolección de datos clínicos pre y perioperatorios División en conjunto de entrenamiento y validación Uso de SMOTE para corregir desequilibrio entre clases (ya que había pocos infectados) Comparación de 7 modelos de IA	Machine Learning tradicional: AdaBoost, Random Forest, Boosted Trees, Gradient Boosting, Extreme Gradient Boosting, GLM, LogitBoost Usaron SMOTE (técnica de balanceo de datos)	Predecir el riesgo de infección en el sitio quirúrgico después de cirugía lumbar Ayudar a identificar pacientes de alto riesgo y mejorar decisiones clínicas	El modelo AdaBoost con SMOTE tuvo el mejor desempeño AUC: 0.906, Recall: 0.9375 Identificó correctamente 15 de 16 pacientes de alto riesgo	Solo se realizó en un hospital (centro único) Se necesita validación externa en otras poblaciones	La IA puede ser útil para predecir SSI tras cirugías de columna Puede ayudar a personalizar la atención y optimizar el manejo perioperatorio

7	Construct validation of machine learning for accurately predicting the risk of postoperative surgical site infection following spine surgery	Corea del sur	2022	5,328 pacientes sometidos a cirugía de columna (espinal), de los cuales 238 desarrollaron infección de sitio quirúrgico (SSI)	Cohorte Estudio retrospectivo multicéntrico	Recolección de 23 variables clínicas preoperatorias de pacientes en 3 hospitales. División 70% entrenamiento y 30% prueba. Ajuste de hiperparámetros. Comparación con regresión logística tradicional	Machine Learning tradicional: XGBoost, LightGBM, Random Forest, Extra Trees, Naive Bayes, Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Multilayer Perceptron (MLP).	Predecir el riesgo de infección del sitio quirúrgico después de cirugía de columna usando solo variables clínicas disponibles antes de la operación	El modelo XGBoost obtuvo el mejor desempeño: AUC = 0.912 También destacaron LightGBM y RF Superaron a la regresión logística tradicional (AUC = 0.738) Alta precisión y sensibilidad para detectar SSI	No se incluyeron variables intra y postoperatorias Necesita validación externa en otros países No se usó NLP ni procesamiento de notas clínicas	Los modelos de ML, especialmente XGBoost, predicen con mayor precisión el riesgo de SSI que los métodos estadísticos tradicionales. Son aplicables en la práctica clínica para toma de decisiones tempranas.
8	Development of machine learning models for the detection of surgical site infections following total hip and knee arthroplasty: a multicenter cohort study	Canadá	2023	22,059 pacientes adultos sometidos a artroplastia total de cadera o rodilla. Se analizaron infecciones del sitio quirúrgico (SSI) profundas, complejas y de espacio órgano.	Cohorte Retrospectivo	Utilizaron datos administrativos estructurados (códigos ICD-10, medicación, procedimientos) y notas clínicas no estructuradas de expedientes electrónicos (EMR). Se dividieron los datos 80% entrenamiento / 20% prueba	XGBoost (Machine Learning supervisado, basado en árboles de decisión ensamblados).	Detectar automáticamente e infecciones postoperatorias en cirugías ortopédicas (cadera y rodilla) utilizando datos clínicos estructurados y texto libre.	AUC-ROC: 0.906- AUC-PR: 0.637- F1-score: 0.788El uso de notas clínicas mejoró sustancialmente el rendimiento del modelo.	Dataset desbalanceado- Solo incluyó notas de enfermería (no notas médicas ni microbiológicas)- No se validó en otros países ni contextos clínicos- Variabilidad en la calidad de las notas clínicas	La IA (XGBoost) puede detectar con alta precisión infecciones en el sitio quirúrgico, automatizando la vigilancia hospitalaria, lo cual reduce la carga de trabajo y mejora la eficiencia en programas de control de infecciones.
9	Development of Multicenter Deep Learning Models for Predicting Surgical Complexity and Surgical Site Infection in Abdominal Wall Reconstruction, a Pilot Study	Estados Unidos	2025	362 pacientes sometidos a reconstrucción de pared abdominal (AWR). Heridas quirúrgicas (infección del sitio quirúrgico, SSI) y complejidad quirúrgica (uso de técnica de separación de componentes, CST)	Estudio piloto retrospectivo	Imágenes preoperatorias de tomografía computarizada (CT) fueron usadas para entrenar modelos de deep learning con arquitectura ResNet-18. Se excluyeron pacientes con imágenes distorsionadas por prótesis o tratamientos previos.	Deep Learning: red neuronal convolucional profunda ResNet-18	Predecir: 1) la complejidad quirúrgica (uso de CST) y 2) la probabilidad de infección postoperatoria (SSI), a partir de imágenes CT antes de la cirugía.	Modelo CST: AUC = 0.569 (bajo rendimiento) Modelo SSI: AUC = 0.898, exactitud = 88.61%, especificidad = 95.65% La predicción de SSIs fue precisa; la predicción de CST fue pobre por su alta subjetividad.	Poca generalización del modelo de CST- Diferencias entre centros quirúrgicos- Tamaño pequeño de cohorte externa- La decisión de usar CST es subjetiva y variable entre cirujanos.	La IA predice con precisión infecciones postoperatorias en AWR usando imágenes preoperatorias. La predicción de decisiones quirúrgicas subjetivas como CST requiere mejores datos o técnicas.

10	Explainable Machine Learning Models for Prediction of Surgical Site Infection After Posterior Lumbar Fusion Surgery Based on Shapley Additive Explanations	China	2025	1016 pacientes consecutivos mayores de 18 años sometidos a cirugía de fusión lumbar posterior por enfermedades degenerativas de la columna (hernia discal lumbar con inestabilidad, espondilolistesis, estenosis espinal). Tipo de herida: infecciones del sitio quirúrgico (SSI)	Cohorte Retrospectivo Unicéntrico	Se recopilaron 44 características clínicas y de laboratorio. Se aplicaron algoritmos de ML para crear modelos predictivos, evaluados con AUC, F1-score, precisión y sensibilidad. Se utilizó validación cruzada k-fold (k=5) y conjunto de prueba independiente	Machine Learning clásico (supervisado) con modelos de ensamble (XGBoost, RF, LightGBM), árboles y vecinos. Explicabilidad con SHAP (Shapley Additive Explanations)	Predecir el riesgo de infección del sitio quirúrgico después de cirugía de columna y explicar qué variables clínicas influyen más en esa predicción	De los 1016 pacientes, 36 (3.54%) desarrollaron SSI. El modelo XGBoost fue el mejor: Precisión: 0.90 Recall: 0.82 Exactitud: 0.99 F1: 0.86 AUC: 0.94	Tamaño pequeño de cohorte, datos de un solo centro, sin validación externa, solo pacientes con cirugía lumbar, sin integración en práctica clínica aún. Sesgos Posible sobreajuste s	Los modelos ML pueden predecir eficazmente el riesgo de SSI y con SHAP pueden explicar de manera transparente las razones de cada predicción. Es útil para la asistencia clínica.
11	Improved Prediction of Surgical-Site Infection After Colorectal Surgery Using Machine Learning	Estados Unidos (UNC Chapel Hill)	2023	275,152 pacientes del NSQIP (2012–2019) con cirugía colorrectal. Heridas quirúrgicas: infección superficial, profunda y de espacio-orgánico.	Cohorte Retrospectivo Validación interna	Los datos de NSQIP fueron divididos en conjuntos de entrenamiento (2012-2018) y prueba (2019). Se probaron 4 modelos: Random Forest, XGBoost, Red Neuronal (NN), y regresión logística. Se evaluó desempeño con AUROC y AUPRC. Validación cruzada y análisis con SHAP para interpretabilidad.	Machine Learning supervisado: Random Forest, XGBoost, Red Neuronal Profunda (Deep Learning), SHAP para explicabilidad.	Mejorar la predicción del riesgo de infección del sitio quirúrgico (SSI) en cirugía colorrectal y superar la precisión de los modelos clásicos como la regresión logística	Mejor modelo: Red Neuronal con AUROC = 0.770 y AUPRC = 0.488. Superó ampliamente a la regresión logística (AUROC = 0.677). Factores más influyentes según SHAP: infección previa al cierre, tiempo operatorio, antibiótico oral, técnica quirúrgica	Sin validación externa Dependencia de variables limitadas del NSQIP No se incluyeron factores clínicos importantes como tipo de antibiótico, tipo de hospital, o técnicas de cierre. No probado aún en entorno clínico real	Los modelos de IA predicen mejor que la regresión logística. Con ajustes y validación externa, podrían incorporarse en sistemas inteligentes (EHR) para personalizar estrategias preventivas.

12	Machine learning application for prediction of surgical site infection after posterior cervical surgery.	China	2024	402 px con cirugía cervical posterior. 44 desarrollaron SSI recolectados de 2018 a 2022	Retrospectivo Cohorte	Se identificaron diferentes variables de interés y se compararon diferentes algoritmos de IA Balanceo de datos con SMOTE-Tomek para corregir el desbalance por baja incidencia de infecciones. Validación cruzada de 10 repeticiones con 10 pliegues Para el análisis metodológico se utilizó R (versión 4.3.1)	RF, GB, ANN, KNN, SVM, LR, DT, NB	Desarrollar y validar modelos predictivos que identifiquen pacientes con alto riesgo de desarrollar infecciones postoperatorias (SSI) tras cirugía de columna cervical posterior, con el objetivo de implementar intervenciones preventivas personalizadas antes de la cirugía.	RF obtuvo la mayor puntuación en AUROC (0.9916), especificidad (0.9890) y precisión (0.9759). GB tuvo la mejor exactitud (0.9535). KNN mostró mayor sensibilidad (0.9958), útil para tamizaje.	Al ser retrospectivo puede involucrar diferentes sesgos, no es multicéntrico y no se incluyen en las variable todos los FR potenciales	Las herramientas de ML permiten predecir con gran precisión el riesgo de infección de herida quirúrgica en pacientes operados de la columna cervical posterior
13	Multicenter Study Protocol: Research on Evaluation and Detection of Surgical Wound Complications with AI-Based Recognition (REDSCAR)	España	2025	168 pacientes post-cirugía abdominal (seguimiento de infecciones en sitio quirúrgico - SSI)	Cohorte Prospectivo Multicéntrico	Se uso la app RedScar© para análisis automático de fotos de Hx Qx y cuestionarios Comparación app vs evaluación presencial RedScar utiliza algoritmos de procesamiento de imágenes con morfología matemática y fuzzy sets, desarrollado en Java SE + Python.	Algoritmos de procesamiento de imágenes con morfología matemática y fuzzy sets, desarrollado en Java SE + Python.	Detectar de forma remota y temprana infecciones quirúrgicas.	sensibilidad 100%, especificidad 83.1%	Necesidad de validar en grandes cohortes y app operador dependiente	Herramienta efectiva para telemonitoreo y detección temprana de complicaciones quirúrgicas. No existe una sección específica de conclusiones en el artículo

14	Preoperative Prediction of Postoperative Infections Using Machine Learning and Electronic Health Record (RME) Data	EU	2024	30,639 px qx (diversos tipos de cx del 2013 al 2019) Infecciones PO: (SSI) (UTI) Sepsis / choque séptico, Neumonía (Todas evaluadas hasta 30 días post-cirugía)	Cohorte Retrospectivo Multicéntrico	Los datos EHR fueron vinculados con ACS-NSQIP. Se dividieron los datos en 70% entrenamiento y 30% prueba. Se seleccionaron de variables con Knockoff filter + LASSO (control de error tipo I). 4 modelos de regresión logística por cada infección. Validación interna. Implementación: R (glmnet, knockoff, pROC).	CNN y Transfer Learning	Estimar el riesgo preoperatorio de infecciones postoperatorias utilizando datos estructurados de registros médicos electrónicos (RME).	AUC en datos de prueba: SSI: 0.73, UTI: 0.76, Sepsis: 0.89, Neumonía: 0.84	Datos de un solo sistema hospitalario → sesgo de generalización. Posible pérdida de datos post-alta. Variabilidad en la calidad de registros EHR. Cambio temporal en comorbilidades → necesidad de recalibración periódica.	Es factible predecir riesgo de infecciones con datos EHR preoperatorios y modelos simples. Requiere validación externa y ajuste por hospital para ser ampliamente implementado.
15	Preliminary Evaluation of a Novel Artificial Intelligence-based Prediction Model for Surgical Site Infection in Colon Cancer	Japón	2022	730 px sometidos a cirugía radical para cáncer de colon en estadio II-III. Px con complicaciones de grado 1-3b de Clavien-Dindo e infecciones es el sitio quirúrgico	Cohorte Retrospectivo unicéntrico	Uso de curvas de Kaplan-Meier para examinar la asociación entre la ISQ y los resultados oncológico. Software de aprendizaje automático Prediction One para predecir la ISQ	Aprendizaje automático Prediction One	Predecir la incidencia de SSI en el cáncer de colon en estadio II-III	Se presentaron ISS en 92 pacientes (12,6%). Entre estos pacientes, 73 presentaron infecciones superficiales y 19 infecciones profundas. AUC: 0,731, derivado mediante análisis estadístico convencional.	1. Es un estudio retrospectivo unicéntrico. 2. No se determinó si los px presentaban trastornos hematológicos o autoinmunes, lo cual podría haber influido en los datos de laboratorio preoperatorios. 3. Los datos no diferenciaron entre infecciones de incisión superficial y profunda. 4. No se realizó validación externa.	El modelo de IA resultó útil para predecir la ISS en pacientes con cáncer de colon en estadio II-III.

16	Using Machine Learning to Predict Surgical Site Infection After Lumbar Spine Surgery	China	2023	Inscritos 4019 px sometidos a cx de fijación interna lumbar en un instituto.	Cohorte Retrospectivo	<p>1. Los datos filtrados se aleatorizaron en los grupos de prueba y verificación.</p> <p>2. En el grupo de prueba, se analizaron variables específicas mediante análisis de regresión logística, análisis de regresión Lasso, máquinas de vectores de soporte y bosque aleatorio.</p>	Técnicas de aprendizaje automático	<p>Analizar factores perioperatorios e identificar los niveles de glucosa en sangre que pueda predecir la aparición de una infección en el sitio quirúrgico después de una cirugía de columna lumbar posterior</p>	<p>Se identificaron 4 FR relacionados con ISS: cambio de Modic, espesor del sebo, Hb y glucosa. Se construyeron curvas ROC y de calibración</p> <p>1. Grupo de prueba: El índice C fue de 0,986 (IC del 95 %: 0,981-0,994).</p> <p>2. Grupo de validación: AUC fue de 0,987. El índice C fue de 0,982 (IC del 95 %: 0,974-0,999)</p>	<p>Puede existir un sesgo de selección y subjetivo debido a la naturaleza retrospectiva del estudio. El modelo podría no ser aplicable a otros centros y requiere verificación externa. Se justifican datos adicionales que podrían mejorar la eficacia diagnóstica del modelo.</p>	<p>Se emplearon análisis de regresión logística y aprendizaje automático para crear un modelo de predicción dinámica para ayudar a los profesionales clínicos a simplificar el monitoreo y la prevención de la Infección del sitio quirúrgico</p>
17	The Use of Artificial Neural Networks for the Prediction of Surgical Site Infection Following TKA	EU	2022	Total: 10,021 px con Artroplastia total de rodilla	Cohorte Retrospectivo	<p>La cantidad total de pacientes, se dividen en dos grupos aleatoriamente, utilizando una proporción estratificada de 80:20, lo que resultó en un conjunto de entrenamiento (8,016) y un conjunto de prueba (2,005)</p>	Uso de 5 métodos supervisados de aprendizaje automático: ANN, SVM, RF y regresión logística penalizada con red elástica.	<p>Demostraron la eficacia de estas técnicas para predecir resultados en pacientes con artroplastia.</p>	<p>El error de puntuación de Brier más bajo se logró con ANN (puntuación de Brier de 0.054). En el conjunto de prueba, las AUC de los 5 modelos oscilaron entre 0.78 y 0.84. La ANN logró la AUC más alta (AUC = 0.84). Las tasas de precisión de los cinco modelos de ML superaron el 94%.</p>	<p>Se demuestra un excelente rendimiento para los 5 modelos candidatos de ML en discriminación, calibración y análisis de la curva de decisión</p>	<p>Los hallazgos del estudio muestran un excelente rendimiento del modelo de aprendizaje automático</p>

18	The Use of Mobile Thermal Imaging and Deep Learning for Prediction of Surgical Site Infection	Rwanda	2021	Imágenes de px PO de Cesárea (530px de las cuales 30 hx estaban infectadas)	Cohorte Retrospectivo	Los datos se utilizaron para desarrollar dos modelos CNN, Se usaron diversos métodos para evitar el sobreajuste y abordar el problema del desequilibrio de clases en la clasificación binaria (infectado y no infectado)	CNN Algoritmos evaluados (Stochastic Gradient Descent (SGD), Adam, and RMSprop) y uso de ResNet 50 model	Predecir la infección del sitio quirúrgico (ISQ) mediante una imagen térmica obtenida con un teléfono inteligente.	1-modelo CNN ingenuo de 6 capas-a precisión mediana de AUC=0,84 con una S del 71 % y una E del 87 %. 2-modelo CNN de aprendizaje por transferencia: precisión mediana de AUC=0,90 con una S del 95 % y una E del 84 %.	Proporción baja de heridas infectadas y no infectadas	Demostración exitosa de un algoritmo de aprendizaje automático para predecir infecciones quirúrgicas utilizando únicamente imágenes térmicas.
19	Use of Convolutional Neural Nets and Transfer Learning for Prediction of Surgical Site Infection from Color Images	China ? Datos obtenidos de artículo de Ruanda	2024	Imágenes de px PO de Cesárea (572 px de las cuales 62 hx estaban infectadas)	Cohorte Retrospectivo	Se utilizaron imágenes de un artículo previamente realizado. Se evaluaron dos CNN, una de ellas con transfer learning. La biblioteca Keras Tensor Flow para competir y entrenar estos modelos CNN en Python. Se realizaron diversos métodos para compensar el desequilibrio y mejorar la sensibilidad	2 CNN: naive CNN model y transfer learning CNN con la arquitectura Resnet 50	Predecir la infección de heridas (cesárea), abordando en particular los problemas de desequilibrio de clases y sobreajuste	1- CNN: mediana de AUC de 0,655, una tasa de verdaderos positivos de 0,75, una especificidad de 0,58 y una precisión de clasificación de 0,86. 2-CNN-con la arquitectura Resnet50: una mediana de AUC de 0,639, una sensibilidad de 0,92, una especificidad de 0,18 y una precisión de clasificación de 0,82.	Grupo de entrenamiento pequeño a comparación de otros estudios	Los modelos CNN mostraron una sensibilidad superior al 80 %, suficiente como herramienta de cribado; sin embargo, la precisión media general del AUC del 64 % es moderada. Debe de considerarse otros aspectos además de solo imágenes

20	Validation and Recalibration of a Model for Predicting Surgical-Site Infection After Pelvic Organ Prolapse Surgery	Estados Unidos	2025	Hx quirúrgicas de mujeres sometidas procedimientos de POP Bases de datos: PHD-recalibración 420.277, 1.6 % resultó en ISS. NSQIP-prueba 62.553, 1.4 % resultó en ISS.	Cohorte Retrospectivo Observacional	Se realiza una recalibración de un modelo previo basado en regresión logística con las bases de datos de PHD y NSQIP Se utilizó el estadístico c y el IC 95 % para evaluar la eficacia	Regresión logística	Predicción de ISQ tras cirugía de POP en una nueva cohorte de mujeres mayores de 40 años.	Discriminación original: c-statistic 0.57 (PHD), 0.59 (NSQIP) (baja). - Recalibración mejoró calibración: intercept -2.18, slope 0.66. - Net benefit: útil para riesgos entre 1-3%. - ICI mejoró de 3.3% a 0.14%. El modelo sobreestimó el riesgo, muy probablemente debido a una menor incidencia de SSI	Uso de datos recolectados de forma retrospectiva	Un modelo desarrollado previamente para predecir la ISQ tras la cirugía de POP demostró una capacidad discriminatoria estable al validarse externamente en las bases de datos PHD y NSQIP. Fue necesaria la recalibración del modelo y se requieren estudios prospectivos.
21	Wound imaging software and digital platform to assist review of surgical wounds using patient smartphones: The development and evaluation of artificial intelligence (WISDOM AI study)	Inglaterra	2024	Fotografías de pacientes de hx quirúrgicas de diferentes etiologías de la biblioteca de Isla Care Ltd. (imágenes de Hx tomadas por personal y por px) Entrenamiento de 37.974 imágenes, Prueba de 3.634 de las cuales se tomó un subconjunto de 355 imágenes para evaluar la fiabilidad intraevaluador	Cohorte Retrospectivo	Aleatorización en Stata versión 18. asignación manual de etiquetas de tono de piel. La revisión por una enfermera especialista se tomó como el "estándar de oro" contra el cual se evaluó el algoritmo de IA. El algoritmo identifica las fotos entre "rutinarias" y "prioritarias"	YOLO	Priorizar las imágenes de heridas quirúrgicas para su revisión por parte del personal clínico en diferentes tonos. Evaluar la confiabilidad intraevaluador del algoritmo de IA	S=del 89% (IC 95%: 87% a 91%), lo que Kappa = 1. Dado que las imágenes se calificaron con una concordancia del 100 %, no se puede calcular un intervalo de confianza del 95 % para Kappa.	No es un algoritmo de IA dinámico, por lo que el proceso de evaluación no cambiará en respuesta a la primera vez	Se supera la sensibilidad objetivo especificada en el protocolo, que era del 85% con un IC 95% de ±3%.

22	Using machine learning to identify risk factors for short-term complications following thumb carpometacarpal arthroplasty	Estados Unidos	2024	Pacientes postoperados de osteoartritis de la articulación carpometacarpiana del pulgar	Cohorte Retrospectiva Cohorte	Los modelos se entrenaron con el 80 % de los datos y se probaron con el 20 % de los datos. El rendimiento del modelo se evaluó utilizando el estadístico Area Under the Receiver Operating Curve (AUC).	ML RF ENet XG Boost NN	Predicción de pacientes en riesgo de presentar complicaciones médicas, complicaciones de heridas y reoperaciones después de la cirugía	El mejor algoritmo para predecir las complicaciones de la herida fue el RF (AUC de $0,64 \pm 0,04$) El algoritmo XGBoost fue el mejor para las complicaciones médicas (AUC: $0,70 \pm 0,06$) y las reintervenciones (AUC: $0,60 \pm 0,07$).	El rendimiento del modelo puede haberse visto afectado por los bajos recuentos absolutos de resultados adversos Tamaño relativamente pequeño de la muestra	Se necesitan ensayos futuros para construir, validar e implementar herramientas de estratificación del riesgo.
23	Multimodal machine learning to predict surgical site infection with healthcare workload impact assessment	Reino Unido	2025	Pacientes adultos sometidos a cirugía gastrointestinal Heridas de sitio operatorio	Cohorte Prospectivo	Recolección de datos de dos estudios (TWIST e INROADE) incluyendo imágenes de heridas y PROMs. Entrenamiento de redes neuronales (MLP, CNN y modelo multimodal) para predecir infección en sitio quirúrgico en 48 horas	MLP CNN	Predecir infecciones del sitio qx en las primeras 48 hrs y reducir la carga de trabajo del personal	AUC multimodal para SSI: 0.762 AUC evaluación clínica: 0.777 (IC 95%: 0.721-0.832) Reducción del tiempo de trabajo del personal: - 82.4% (de 51.5 h a 9.1 h por cada 100 pacientes) Disminución de revisiones clínicas: de 1545 a 272 respuestas	Tasa de eventos bajos Imágenes heterogéneas en calidad/resolución	La red neuronal multimodal predijo infecciones con precisión comparable a la evaluación clínica, reduciendo significativamente la carga de trabajo. Es viable para su uso en monitoreo remoto postoperatorio.

Anexo 5: Tabla de evaluación a la adherencia metodológica con TRIPOD+AI heridas crónicas

	Automated Prediction of Photographic Wound Assessment Tool	Automatic Classification of Diabetic Foot Ulcer Images - A Transfer Learning Approach to Detect Wound Necrotization	An artificial intelligence-enabled smartphone app for real-time pressure injury assessment	Chronic Ulcers Healing Prediction through ML Approaches (Preliminary DFU Case Study)	Clinical validation of an AI-enabled wound imaging mobile application in DFUs	DFUCare: deep learning platform for diabetic foot ulcers	Diabetic foot ulcer classification assessment	Guided Conditional Diffusion Classifier (CoDF) for DFU infection prediction	ML algorithm to evaluate risk factors of DFU severity	ML Models for Predicting Hard-to-Heal DFUs in Chinese Population	Prediction of Foot Ulcers Using AI (Cairn University Hospital)	Prediction model of pressure injury occurrences in diabetic patients during ICU hospitalization - XGBoost model (intolerable with SHAP)	Region-Based Semi-Two-Stream Convolutional Neural Networks for Pressure Ulcer Recognition	Supervised machine learning-based prediction for in-hospital pressure injuries	Survival Prediction in Diabetic Foot Ulcers	Visual classification of pressure injury stages for nurses: a deep learning model with modern CNNs	A feature explainability-based deep learning technique for diabetic foot ulcer identification	Machine learning-driven discovery of novel therapeutic targets in diabetic foot ulcers	Development and application of an intelligent pressure injury assessment system using AI image recognition	Accelerating Surgical Site Infection Abstraction With a Semi-automated Machine Learning Approach
1. Título identifica modelo, población y outcome	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1
2. Abstract adecuado (TRIPOD-AI)	1	1	1	1	0	0	1	1	1	0	1	1	1	1	1	0	0	1	1	0
3a. Contexto y justificación	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
3b. Población objetivo e intención de uso	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
3c. Considera desigualdades en salud	0	0	0	1	1	1	1	1	1	1	0	0	1	0	1	0	1	1	0	1
4. Objetivos claros	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
5a. Fuentes de datos (desarrollo/evaluación)	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1
5b. Fechas de recolección	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
6a. Entorno y centros	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
6b. Criterios de elegibilidad	1	0	1	0	1	1	0	0	1	1	1	1	0	0	1	0	0	1	0	0
6c. Tratamientos y manejo relacionados con el outcome	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
7. Preparación/calidad de datos	1	1	1	1	0	1	1	1	0	0	1	0	1	1	1	1	1	1	1	1
8a. Definición del outcome y horizonte temporal	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
8b. Perfil de evaluadores del outcome	1	1	0	0	1	1	0	1	0	1	1	1	1	0	1	0	1	0	1	1
8c. Cegamiento en evaluación del outcome	0	0	0	0	1	1	0	0	0	0	0	0	1	0	1	0	1	0	0	1
9a. Predictores iniciales justificados	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
9b. Definición y medición de predictores	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
9c. Perfil de evaluadores de predictores	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
10. Tamaño muestral justificado	0	0	1	0	1	0	0	0	1	0	1	1	0	0	1	0	1	1	0	0
11. Manejo de datos faltantes	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12a. Uso de datos y particiones	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
12b. Manejo de predictores en el análisis	1	1	1	1	0	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1
12c. Tipo de modelo y validación interna	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1
12d. Heterogeneidad entre clusters	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12e. Medidas de desempeño	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
12f. Recalibración del modelo	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
12g. Cómo obtener predicciones	1	1	1	1	0	1	1	1	1	0	1	0	1	0	1	1	1	1	1	1
13. Manejo de desbalance de clases	0	0	1	0	1	0	0	1	0	1	0	0	1	0	1	0	1	0	0	0
14. Métodos de fairness	0	0	0	1	1	1	1	0	1	1	0	0	0	1	1	0	1	0	1	0
15. Output del modelo y umbrales	1	0	1	0	1	1	0	1	1	0	1	1	1	1	0	0	0	0	0	1
16. Diferencias entre datasets de desarrollo y validación	1	0	0	0	0	1	0	0	0	0	0	0	0	1	1	0	0	0	0	0
17. Aprobación ética	1	0	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	0
18a. Financiamiento	1	1	1	0	1	1	0	1	1	0	0	1	1	1	1	1	1	1	1	1
18b. Conflictos de interés	1	0	1	0	1	0	0	1	1	0	0	1	1	1	0	1	1	1	0	0
18c. Protocolo accesible	0	0	0	1	1	0	0	1	0	0	0	0	0	0	1	0	0	0	0	1
18d. Registro del estudio	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
18e. Disponibilidad de datos	1	0	0	0	1	1	0	0	0	0	0	0	1	0	0	0	0	0	0	1
18f. Disponibilidad de código	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
19. Participación de pacientes/público	0	1	0	1	0	1	0	0	0	0	0	0	0	1	0	0	1	1	0	1
20a. Flujo de participantes	0	0	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0
20b. Características reportadas	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
20c. Comparación con dataset de desarrollo	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
21. N participantes y eventos	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1
22. Especificación completa del modelo	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
23a. Desempeño con IC	0	0	1	0	1	0	0	0	1	1	0	0	0	1	1	0	0	0	0	0
23b. Heterogeneidad del desempeño	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
24. Resultados de recalibración	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
25. Interpretación global	1	0	1	1	0	1	0	1	1	0	0	1	0	1	0	0	1	1	0	0
26. Limitaciones discutidas	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1
27a. Implementación: datos faltantes	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
27b. Implementación: experiencia del usuario	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
27c. Próximos pasos y aplicabilidad	1	0	0	1	0	1	1	1	1	1	0	1	0	1	0	0	1	0	1	1
TOTAL	32	26	34	32	31	37	27	32	33	27	28	31	31	36	32	32	34	29	29	33
Porcentaje	61.5	50.0	65.4	61.5	59.6	71.2	51.9	61.5	63.5	51.9	53.8	59.6	59.6	69.2	61.5	61.5	65.4	51.9	51.9	63.5

Anexo 6: Tabla de evaluación a la adherencia metodológica con TRIPOD+AI con heridas quirúrgicas

	Applying Machine Learning Across Sites: External Validation of a Surgical Site Infection Detection Algorithm (Zhu 2021)	Artificial Intelligence-Based Multimodal Risk Assessment Model for Surgical Site Infection (AMRAMS): Development and Validation Study (Chen 2020)	Can machine learning models improve the prediction of surgical site infection in abdominal surgery than traditional statistical models? (Piebpien 2024)	Construct validation of machine learning for accurately predicting the risk of postoperative surgical site infection following spine surgery (Zhang 2024)	ML for SSI after THA/TKA (Wu 2023)	Development of Multicenter Deep Learning Models for Predicting Surgical Site Infection in Abdominal Wall Reconstruction, a Pilot Study (Lorenz 2025)	Explainable Machine Learning Models for Prediction of Surgical Site Infection After Posterior Lumbar Fusion Surgery Based on Shapley Additive Explanations (Wang 2025)	A machine learning model to predict surgical site infection after surgery of lower extremity fractures	Accelerating Surgical Site Infection Abstraction With a Semi-automated Machine-learning Approach	Improved Prediction of Surgical-Site Infection After Colorectal Surgery Using Machine Learning	Machine learning application for prediction of surgical site infection after posterior cervical surgery	Multicenter Study Protocol: Research on Evaluation and Detection of Surgical Wound Complications with AI-Based Recognition (REDS-CAR-Trial)	Preliminary Evaluation of a Novel Artificial Intelligence-based Prediction Model for Surgical Site Infection in Colon Cancer	Preoperative Prediction of Postoperative Infections Using Machine Learning and Electronic Health Record Data	The Use of Artificial Neural Networks for the Prediction of Surgical Site Infection Following TKA
1. Título identifica modelo, población y outcome	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
2. Abstract adecuado	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
3a. Contexto y justificación	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
3b. Población objetivo definida	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
3c. Considera desigualdades en salud	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4. Objetivos claros	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
5a. Fuente de datos descrita	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
5b. Fechas de recolección	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
6a. Entorno/centros	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
6b. Criterios de elegibilidad	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
6c. Tratamientos descritos	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7. Preparación/calidad de datos	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
8a. Outcome definido claramente	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
8b. Evaluadores del outcome capacitados	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8c. Cegamiento en evaluación del outcome	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9a. Predictores justificados	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
9b. Predictores definidos	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
9c. Evaluadores de predictores	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10. Tamaño muestral justificado	0	1	1	1	1	0	1	1	0	0	0	0	0	0	0
11. Manejo de datos faltantes	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
12a. Uso de datos (train/test, validación)	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1
12b. Manejo de predictores (transformación, estandarización)	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1
12c. Tipo de modelo, tuning y validación interna	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1
12d. Heterogeneidad entre clusters	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12e. Medidas de desempeño reportadas	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
12f. Actualización/recalibración	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0
12g. Cómo se calcularon predicciones	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0
13. Manejo de desbalance de clases	1	1	1	1	1	1	1	1	0	1	0	0	0	0	0
14. Métodos de fairness	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15. Output del modelo especificado	1	1	1	1	1	1	1	1	1	1	0	0	1	0	0
16. Diferencias entre datos desarrollo/evaluación	1	1	0	0	1	0	0	0	0	1	0	0	1	1	1
17. Aprobación ética	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
18a. Financiamiento	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
18b. Conflictos de interés	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
18c. Protocolo accesible	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
18d. Registro del estudio	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
18e. Disponibilidad de datos	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0
18f. Disponibilidad de código	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
19. Involucramiento de pacientes/público	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
20a. Flujo de participantes descrito	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1
20b. Características reportadas	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1
20c. Comparación desarrollo vs validación	1	1	0	0	1	0	0	0	0	0	0	0	0	0	0
21. N pacientes y eventos en cada análisis	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0
22. Especificación completa del modelo para replicación	0	0	1	1	0	1	1	1	1	0	0	0	0	0	0
23a. Medidas de desempeño con IC	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1
23b. Heterogeneidad de desempeño entre clusters	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0
24. Resultados de recalibración (si hubo)	0	0	1	1	0	1	1	1	1	0	0	0	0	0	0
25. Interpretación global	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1
26. Limitaciones discutidas	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1
27a. Manejo de datos faltantes en implementación	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0
27b. Nivel de experiencia requerida de usuarios	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
27c. Próximos pasos y aplicabilidad futura	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
TOTAL	35	35	36	34	35	33	34	36	33	27	29	16	27	28	27
Porcentaje adherencia	67,3	67,3	69,2	65,4	67,3	63,5	65,4	69,2	63,5	51,9	55,8	30,8	51,9	53,8	51,9