

Aplicaciones de la inteligencia artificial en la educación quirúrgica de posgrado: revisión sistemática

Applications of artificial intelligence in postgraduate surgical education: a systematic review

Johanna Yimabel Pazmiño Figueroa¹ [0009-0004-8368-159X] María Eugenia Moreta Segura² [0009-0005-0568-7651]
Wellington Isaac Maliza Cruz³ [0009-0005-1426-583X]

Universidad Estatal de Milagro. Facultad de Posgrados. Escuela de Educación. Maestría en Inteligencia Artificial. Cdla. Universitaria “Dr. Rómulo Minchala Murillo” – km. 1.5 vía Milagro – Virgen de Fátima. CP 091050. Milagro - Guayas. Ecuador.

yimabel.pazmino@educacion.gob.ec eugenia.moreta@educacion.gob.ec wmaliza@utb.edu.ec

CITA EN APA:

Pazmiño Figueroa, J. Y., Moreta Segura, M. E., & Maliza Cruz, W. I. (2026). APLICACIONES DE LA INTELIGENCIA ARTIFICIAL EN LA EDUCACIÓN QUIRÚRGICA DE POSGRADO: REVISIÓN SISTEMÁTICA. *Tesla Revista Científica*, 6(1), e670.

<https://doi.org/10.55204/trc.v6i1.e670>

Recibido: 2026-06-20

Aceptado: 2026-06-22

Publicado: 2026-06-22

TESLA
Revista Científica
ISSN: 2953-4275



Los contenidos de este artículo están bajo una licencia de Creative Commons Attribution 4.0 International (CC BY 4.0)

Los autores conservan los derechos morales y patrimoniales de sus obras.

The contents of this article are under a Creative Commons Attribution 4.0 International (CC BY 4.0) license. The authors retain the moral and patrimonial rights of their works.

Resumen

Objetivo: Analizar la evidencia científica disponible sobre el uso de la inteligencia artificial (IA) en la educación quirúrgica de posgrado.

Metodología: Se realizó una revisión sistemática siguiendo estándares metodológicos internacionales. La búsqueda se efectuó en PubMed, Scopus, Web of Science y Embase, incluyendo estudios publicados entre 2015 y 2025. Se incluyeron ensayos clínicos aleatorizados, estudios experimentales y de validación tecnológica enfocados en médicos en formación de posgrado.

Resultados: Se seleccionaron 29 estudios. La evidencia muestra que la IA mejora significativamente el desempeño técnico quirúrgico, reduce errores y optimiza el tiempo de aprendizaje mediante retroalimentación automatizada y evaluación objetiva. Además, permite la personalización del entrenamiento y contribuye a la estandarización de competencias. Sin embargo, se identificó heterogeneidad metodológica, tamaños muestrales limitados y escasa validación externa.

Conclusiones: La inteligencia artificial representa una herramienta innovadora con alto potencial para mejorar la formación quirúrgica de posgrado. No obstante, su implementación debe ser progresiva, complementaria a la enseñanza tradicional y basada en evidencia sólida.

Palabras Clave: Inteligencia artificial; educación quirúrgica; formación médica; simulación; aprendizaje automático.

Abstract

Objective: To analyze the available scientific evidence on the use of artificial intelligence (AI) in postgraduate surgical education.

Methodology: A systematic review was conducted following international methodological standards. The search was carried out in PubMed, Scopus, Web of Science, and Embase, including studies published between 2015 and 2025. Randomized clinical trials, experimental studies, and technological validation studies focused on physicians undergoing postgraduate training were included.

Results: A total of 29 studies were selected. The evidence shows that AI significantly improves surgical technical performance, reduces errors, and optimizes learning time through automated feedback and objective assessment. In addition, it enables the personalization of training and contributes to the standardization of competencies. However, methodological heterogeneity, limited sample sizes, and scarce external validation were identified.

Conclusions: Artificial intelligence represents an innovative tool with high potential to improve postgraduate surgical training. Nevertheless, its implementation should be progressive, complementary to traditional teaching, and based on solid evidence.

Keywords: Artificial intelligence; surgical education; medical training; simulation; machine learning.

1. INTRODUCCIÓN

La preparación quirúrgica durante la formación de posgrado constituye una etapa esencial en el desarrollo de competencias médicas avanzadas, porque integra habilidades técnicas, razonamiento clínico y toma de decisiones en contextos complejos. A pesar de su importancia, los modelos tradicionales basados en la práctica supervisada encuentran limitaciones, entre ellas el tiempo disponible para el entrenamiento, seguridad del paciente y la exposición clínica (1). Por ello, resulta necesario incorporar metodologías educativas innovadoras que fortalezcan y permitan una preparación eficaz, estandarizada y sobre todo segura.

Dentro de este contexto, la inteligencia artificial ha adquirido un papel cada vez más relevante como herramienta de apoyo en la educación quirúrgica de posgrado. Su utilidad se relaciona con la posibilidad de realizar evaluaciones más objetivas del desempeño, generar retroalimentación automatizada y potenciar experiencias de aprendizaje mediante simulaciones avanzadas (4). Además, su empleo permite adaptar la enseñanza a las necesidades individuales del residente, analizar con mayor precisión las habilidades técnicas e identificar de forma temprana posibles dificultades durante la formación, aspectos que pueden favorecer una mejora progresiva de las curvas de aprendizaje (2).

Sin embargo, aunque sus beneficios son prometedores, la evidencia disponible muestra limitaciones metodológicas y desafíos vinculados con su validación en entornos clínicos reales. En consecuencia, es fundamental analizar críticamente su impacto para orientar una incorporación adecuada y responsable en los programas de formación médica. En este contexto, la inteligencia artificial cumple un papel cada vez más importante como herramienta de apoyo a la educación quirúrgica. Sus múltiples usos se relacionan con la posibilidad de generar retroalimentación, potenciar experiencias de aprendizaje mediante procesos de simulaciones y la oportunidad de realizar evaluaciones con mayor grado de objetividad (3).

Por otro lado, su uso adapta las necesidades individuales de cada residente, analizar su desempeño y luego adquirir retroalimentación con un mayor grado de precisión. Sin embargo, aunque sus beneficios son prometedores, la evidencia disponible aún muestra limitaciones metodológicas y desafíos vinculados con su comparabilidad en ambientes clínicos reales. Por lo tanto, es fundamental analizar críticamente su impacto para orientar una incorporación adecuada y responsable en los programas de formación médica (3,4).

La educación quirúrgica en el nivel de posgrado se enfrenta actualmente a desafíos importantes tanto en su organización como en sus métodos de enseñanza. Entre los principales problemas se encuentran la desigualdad en la exposición de los residentes a casos clínicos, el tiempo reducido para la práctica, la necesidad permanente de proteger la seguridad del paciente y la falta de herramientas objetivas y uniformes para evaluar el desempeño quirúrgico. En este contexto, la incorporación de tecnologías innovadoras se

vuelve cada vez más necesaria, ya que puede contribuir a mejorar la calidad del aprendizaje y a optimizar la formación médica especializada (2).

El aprendizaje quirúrgico a nivel de posgrado se enfrenta actualmente a desafíos en su organización debido a su fuerte carga asistencial y el número de cirugías disponibles para los residentes, el equilibrio entre el tiempo quirúrgico y el valor asistencial del residente es difícil de alcanzar, por lo que, es necesario encontrar herramientas que permitan simulaciones que permitan la exposición de los residentes a casos clínicos. Sin embargo, entre los principales problemas se encuentran el tiempo reducido para la práctica, la necesidad permanente de proteger la seguridad del paciente y la falta de herramientas objetivas (2).

Frente a estas dificultades, la inteligencia artificial surge como una opción con alto potencial para transformar la enseñanza quirúrgica. Su aplicación permite desarrollar sistemas capaces de analizar grandes cantidades de datos clínicos y educativos, evaluar con mayor precisión las habilidades técnicas y entregar retroalimentación rápida, personalizada y ajustada al rendimiento de cada residente. A diferencia de los modelos tradicionales, que suelen depender principalmente del criterio del docente o tutor, la IA facilita el uso de parámetros cuantificables, repetibles y comparables. Esto favorece una evaluación más precisa y fortalece un proceso formativo centrado en la adquisición progresiva de competencias (3).

La aplicación de la inteligencia artificial permite desarrollar sistemas capaces de analizar, evaluar con mayor precisión habilidades técnicas y generar retroalimentación ajustada a cada residente. A diferencia de los modelos tradicionales, que suelen depender principalmente del criterio del docente o tutor, la IA facilita el uso de parámetros cuantificables, repetibles y comparables. Esto favorece una evaluación más precisa y fortalece un proceso formativo centrado en la adquisición progresiva de competencias (3).

De igual manera, el uso de inteligencia artificial en simuladores quirúrgicos y plataformas digitales de aprendizaje permite crear espacios de entrenamiento más seguros y controlados. En estos entornos, los residentes pueden practicar de forma continua, corregir errores y perfeccionar sus destrezas sin poner en riesgo al paciente. Este aporte resulta especialmente relevante en la actualidad, debido a que la disponibilidad limitada de casos clínicos complejos y la necesidad de utilizar mejor los recursos disponibles pueden restringir las oportunidades de aprendizaje mediante los métodos quirúrgicos convencionales (4).

2. METODOLOGÍA O MATERIALES Y MÉTODOS

Se realizó una revisión sistemática con el propósito de reunir, examinar y sintetizar la evidencia científica relacionada con el uso de la inteligencia artificial en la formación quirúrgica de posgrado. Para cumplir este objetivo, se siguió un proceso metodológico ordenado que comprendió la identificación de estudios, selección de publicaciones, análisis crítico de la calidad de la evidencia y posterior integración de los hallazgos. La búsqueda de información se planteó de manera amplia y se ejecutó en bases de datos

reconocidas en los campos biomédico y tecnológico, como MEDLINE/PubMed, Scopus, Web of Science y Embase. Con la finalidad de obtener una revisión más completa y disminuir la posibilidad de excluir estudios relevantes, no se aplicaron restricciones por idioma ni por especialidad quirúrgica.

La estrategia de búsqueda bibliográfica se elaboró a partir de la combinación de términos controlados y palabras clave relacionadas con la inteligencia artificial y la educación quirúrgica. Se incluyeron conceptos como “artificial intelligence”, “machine learning”, “deep learning”, “surgical education”, “surgical training”, “residency”, “simulation”, “feedback”, “intelligent tutor” y “virtual reality”. Estos términos fueron combinados mediante operadores booleanos como AND y OR, lo que permitió ampliar la recuperación de artículos pertinentes y, al mismo tiempo, hacer más específica la selección de resultados.

Para la construcción de la estrategia se tomaron como referencia revisiones sistemáticas previas de adecuada calidad metodológica, como las realizadas por Kirubarajan et al. y Alaa et al., en las que se aplicaron búsquedas extensas en bases de datos interdisciplinarias. Para la selección de los artículos, se definieron criterios de inclusión dirigidos a identificar estudios originales publicados entre 2015 y 2025, enfocados en aplicaciones de inteligencia artificial dentro de la formación quirúrgica de médicos residentes o profesionales en etapa de posgrado.

Se consideraron investigaciones con distintos diseños metodológicos, incluyendo ensayos clínicos aleatorizados, estudios cuasi-experimentales, cohortes prospectivas y trabajos de validación tecnológica con aplicación educativa. Por el contrario, se excluyeron documentos sin datos empíricos, como revisiones narrativas, editoriales y cartas al editor. También fueron descartados los estudios realizados en estudiantes de pregrado, aquellos que no evaluaban directamente intervenciones educativas basadas en IA y las investigaciones centradas únicamente en modelos animales o simulaciones sin una finalidad formativa concreta dentro del entrenamiento quirúrgico.

La selección de los estudios se llevó a cabo mediante un proceso ordenado y progresivo, dividido en dos fases principales. En primer lugar, se examinaron los títulos y resúmenes de los registros obtenidos durante la búsqueda, lo que permitió eliminar artículos duplicados y excluir aquellos que no se relacionaban directamente con el propósito de la revisión. Luego, las publicaciones consideradas potencialmente relevantes fueron evaluadas en texto completo, aplicando los criterios de elegibilidad establecidos con anterioridad para determinar su inclusión definitiva. De un total aproximado de 500 registros identificados inicialmente, tras la depuración, revisión y evaluación de pertinencia, se seleccionaron finalmente 28 estudios para la síntesis cualitativa.

La calidad metodológica de los estudios incluidos fue evaluada mediante herramientas validadas, tomando en cuenta elementos como el riesgo de sesgo, la consistencia de los resultados y la aplicabilidad de los hallazgos en contextos clínicos y educativos. En conjunto, la evidencia disponible mostró un nivel

moderado; sin embargo, también se observaron debilidades recurrentes, principalmente asociadas al tamaño limitado de las muestras y a la variedad de intervenciones analizadas. Esta diversidad metodológica puede explicarse por la amplitud de los criterios de inclusión y exclusión empleados en la revisión.

3. DESARROLLO

La preparación quirúrgica de posgrado es un proceso en el que el médico en formación debe adquirir diferentes destrezas en el ámbito quirúrgico, razonamiento clínico y capacidad de toma de decisiones dentro de un quirófano. En este contexto, la seguridad del paciente y la calidad de abordaje del aprendizaje deben avanzar de manera conjunta, ya que deben ir de la mano como elementos complementarios dentro del entrenamiento médico especializado (3,5).

Desde una mirada pedagógica actual, la incorporación de la inteligencia artificial en este ámbito se fundamenta principalmente en dos ejes: la formación orientada al desarrollo de competencias y la retroalimentación con finalidad formativa (3). En este enfoque, el avance del residente no se determina solo por el tiempo de permanencia en el programa académico, sino por la evidencia progresiva, objetiva y comprobable de los aprendizajes y habilidades adquiridas. Este modelo ha sido ampliamente abordado en el campo de la educación médica, sobre todo en relación con su conceptualización, integración curricular y formas de evaluación de los resultados alcanzados (5).

En este caso, el tener la oportunidad de una retroalimentación objetiva, clara, específica y dirigida y adaptada al residente. Desde el enfoque de la teoría del feedback, sus efectos sobre el proceso de aprendizaje no son uniformes, ya que dependen de cómo se comunica la información, del contexto educativo en el que se aplica y de la interpretación que realiza el aprendiz. En este sentido, la retroalimentación automatizada mediante inteligencia artificial no debería limitarse a emitir juicios técnicos, sino integrarse dentro de una estrategia pedagógica bien estructurada. Esto favorece que las orientaciones generadas sean relevantes, aplicables y acordes con las competencias previstas en el entrenamiento quirúrgico (6).

Un aspecto importante dentro de la formación quirúrgica contemporánea es la práctica deliberada, la cual se comprende como un proceso organizado en el que el desarrollo del desempeño experto se logra mediante la repetición guiada, la definición de objetivos específicos, la supervisión continua y la corrección gradual de los errores. Este enfoque respalda el uso de simuladores y de instrumentos de evaluación que permitan valorar el rendimiento de forma más objetiva (4).

En el ámbito quirúrgico, herramientas como el Objective Structured Assessment of Technical Skill (OSATS) se han planteado como métodos estructurados, válidos y confiables para evaluar las competencias técnicas de los residentes. En consecuencia, constituyen una base relevante para la creación de métricas automatizadas destinadas a valorar el desempeño en cirugía (7).

A partir de estos fundamentos, el soporte pedagógico del proyecto puede estructurarse en torno a tres requerimientos esenciales: impulsar la práctica deliberada, reforzar la evaluación sistemática y asegurar una retroalimentación formativa adecuada. En este sentido, las tecnologías que permiten registrar, procesar y analizar el desempeño de manera continua, objetiva y escalable pueden contribuir al perfeccionamiento del aprendizaje quirúrgico, siempre que sean incorporadas como complemento de la supervisión docente y de los modelos educativos centrados en competencias (3,4).

Por consiguiente, el análisis de la inteligencia artificial no debe enfocarse en considerarla un reemplazo del profesor, sino como una herramienta de apoyo capaz de facilitar una medición más precisa del avance del residente, reducir la variabilidad en los procesos de entrenamiento y promover una enseñanza más homogénea y coherente.

3.1.TECNOLOGÍAS

La aplicación de la inteligencia artificial en la educación quirúrgica de posgrado puede abordarse desde diferentes líneas tecnológicas, las cuales aportan de manera diferenciada al fortalecimiento del aprendiz y la evaluación de residentes. Una de las áreas más relevantes corresponde al uso de aprendizaje automático y aprendizaje profundo, sobre todo cuando estas herramientas se integran con sistemas de visión por computador.

Esta combinación permite procesar y analizar videos obtenidos durante procedimientos quirúrgicos, A través de este recurso es posible reconocer las distintas etapas del procedimiento, identificar el instrumental empleado, seguir la trayectoria de los movimientos y evaluar la precisión de los gestos técnicos. Así, la actividad quirúrgica puede convertirse en datos medibles que contribuyen a valorar de forma más objetiva el rendimiento del residente (8).

En el ámbito de la enseñanza quirúrgica, esta posibilidad resulta especialmente relevante, debido a que el video ya se utiliza como apoyo para el aprendizaje, la supervisión clínica y la retroalimentación docente. Además, proporciona una base más objetiva para estudiar la ejecución técnica. En este contexto, la inteligencia artificial puede reducir la dependencia exclusiva del juicio de expertos, un proceso que generalmente requiere tiempo y puede variar según los criterios de cada evaluador (2).

Dentro de las aplicaciones más destacadas se encuentra el uso del aprendizaje automático y el aprendizaje profundo, particularmente cuando estas tecnologías se combinan con la visión por computador. Esta asociación permite examinar videos obtenidos durante actos quirúrgicos, identificando patrones y elementos relevantes de la ejecución técnica.

Asimismo, los sistemas de tutoría inteligente y las plataformas que ofrecen retroalimentación automatizada representan una línea importante de desarrollo. Estas herramientas emplean métricas de

rendimiento junto con algoritmos o criterios de decisión para supervisar la evolución del residente a lo largo de su formación. Más allá de medir o predecir el grado de destreza alcanzado, su propósito es brindar comentarios oportunos, homogéneos y acordes con las necesidades individuales del aprendiz. Para que dicho proceso tenga valor educativo, la retroalimentación debe relacionarse directamente con los métodos de evaluación, los objetivos de aprendizaje y las competencias contempladas en el programa quirúrgico (3,6).

Por otra parte, los modelos de lenguaje de gran escala y el procesamiento del lenguaje natural incorporan nuevas posibilidades en la formación quirúrgica de posgrado. Estas herramientas pueden apoyar la elaboración de retroalimentación narrativa, la explicación del razonamiento clínico y la creación de casos, preguntas o escenarios simulados para el aprendizaje. La evidencia disponible se ha enfocado principalmente en su capacidad para detectar errores y producir comentarios pertinentes en tareas escritas, lo que favorece el desarrollo de habilidades cognitivas, pensamiento crítico y autorreflexión sobre el desempeño del residente (2).

Para que estas aplicaciones puedan avanzar tanto en la investigación como en la enseñanza quirúrgica, resulta indispensable contar con conjuntos de datos adecuados. En este contexto, JIGSAWS constituye uno de los recursos más utilizados para estudiar la actividad quirúrgica y valorar las destrezas técnicas en cirugía robótica. Este dataset fue desarrollado a partir de registros generados mediante un sistema robótico e incluyó a ocho cirujanos con diversos niveles de experiencia, quienes ejecutaron varias repeticiones de tareas fundamentales, entre ellas sutura, anudado y paso de aguja (9,10).

Además, combina videos sincronizados con datos cinemáticos, lo que permite analizar el desempeño quirúrgico desde perspectivas complementarias (9,10). Su valor en la educación quirúrgica se debe, principalmente, a que las tareas evaluadas forman parte del entrenamiento técnico básico y a que ofrece un marco común para comparar el desempeño de distintos algoritmos.

3.2. JIGSAWS

Las investigaciones basadas en JIGSAWS también han puesto en evidencia una dificultad relevante: la limitada disponibilidad de datasets abiertos y la falta de métodos de validación plenamente estandarizados. La existencia de una referencia compartida favorece una comparación más precisa entre métodos tecnológicos y permite valorar de mejor manera su evolución. Algunos estudios han informado porcentajes cercanos al 80 % en el reconocimiento de gestos quirúrgicos, dependiendo del esquema de validación utilizado, lo cual refleja avances importantes en esta área. A pesar de ello, todavía existen limitaciones para trasladar estos resultados a contextos más diversos, especialmente cuando se consideran diferencias entre cirujanos en cuanto a experiencia, estilo técnico y manera de realizar los procedimientos (8).

Los sistemas de tutoría inteligente y las plataformas de aprendizaje adaptativo representan otra línea relevante dentro de la aplicación de la inteligencia artificial en la educación quirúrgica. Estas herramientas se basan en el análisis del desempeño del residente para adaptar los contenidos, modificar el nivel de dificultad de las actividades y regular tanto la frecuencia como el tipo de retroalimentación proporcionada.

En este sentido, Goldenberg et al. evidenciaron que la enseñanza apoyada por IA puede disminuir la dependencia de la intervención constante del instructor, sin afectar negativamente los resultados formativos e, incluso, con la posibilidad de mejorarlos (12). Este modelo favorece un uso más racional de los recursos educativos y promueve estrategias de enseñanza más individualizadas, centradas en las necesidades y el progreso del estudiante.

3.3.TUTORÍAS

Los sistemas de tutoría inteligente y las plataformas de aprendizaje adaptativo representan otra línea relevante dentro de la aplicación de la inteligencia artificial en la educación quirúrgica. Se basan en el análisis de desempeño para adaptar los contenidos, modificar el nivel de dificultad de las actividades y regular tanto la frecuencia como el tipo de retroalimentación proporcionada. Desde esta perspectiva, Goldenberg et al. señalaron que el apoyo de la inteligencia artificial en la enseñanza puede disminuir la dependencia de la supervisión permanente del docente, sin generar efectos negativos en los resultados educativos y, en algunos casos, contribuyendo a su mejora (12).

Este tipo de modelo facilita una gestión más adecuada de los recursos formativos y promueve estrategias de aprendizaje más individualizadas, centradas en el ritmo, las necesidades y el progreso del estudiante. Las aplicaciones de la inteligencia artificial en la formación quirúrgica de posgrado pueden organizarse de acuerdo con la función que desempeñan dentro del proceso educativo y con el tipo de tecnología utilizada. Su aporte no se restringe a una única herramienta, sino que incluye recursos dirigidos al entrenamiento práctico, la valoración del desempeño y la retroalimentación del aprendizaje.

Una de las líneas con mayor desarrollo corresponde a la simulación quirúrgica apoyada por inteligencia artificial. En este campo, los simuladores basados en realidad virtual o aumentada incorporan algoritmos de aprendizaje automático capaces de registrar, procesar e interpretar los movimientos realizados por el residente durante la práctica. Con base en estos datos, se pueden detectar errores técnicos, identificar patrones de ejecución y proporcionar retroalimentación inmediata, con la finalidad de promover la corrección progresiva de fallos y optimizar el rendimiento quirúrgico (4).

Otra aplicación importante se relaciona con la evaluación objetiva del desempeño y la producción automatizada de retroalimentación. Diversos estudios han empleado modelos de aprendizaje profundo para analizar videos quirúrgicos y generar mediciones cuantitativas del rendimiento, obteniendo resultados más reproducibles y menos influenciados por la subjetividad del evaluador (Chen et al., Hung et al.) (23,24).

De forma similar, las herramientas de visión por computador han demostrado utilidad en el análisis de habilidades técnicas con elevados niveles de precisión y una concordancia significativa con las valoraciones realizadas por expertos humanos (Ward et al.) (11). En conjunto, revisiones recientes indican que estas tecnologías pueden contribuir a una retroalimentación más objetiva, estructurada y pertinente, fortaleciendo el proceso de entrenamiento quirúrgico de los residentes (Kirubarajan et al., Goldenberg et al.) (12,13).

Por último, la inteligencia artificial generativa y las técnicas de procesamiento del lenguaje natural constituyen una línea emergente en la educación quirúrgica de posgrado. Aunque la evidencia disponible aún es limitada, estas herramientas muestran potencial para elaborar casos clínicos simulados, facilitar procesos de tutoría automatizada y apoyar la evaluación de competencias cognitivas relacionadas con el razonamiento clínico y la toma de decisiones. Estudios recientes sugieren que su incorporación podría complementar las estrategias educativas tradicionales; sin embargo, su uso en la formación médica exige procesos de validación rigurosos y una reflexión cuidadosa sobre sus implicaciones éticas, pedagógicas y profesionales.

Estudio	Número de participantes y país	Diseño	Seguimiento	Objetivos	Resultados
Wu et al. (2024) (14)	n=22 cirujanos novatos, China	Ensayo clínico aleatorizado	5 procedimientos	Evaluar coaching con IA en laparoscopia	Incremento significativo en la calidad del desempeño quirúrgico, con mayor seguridad intraoperatoria y reducción de eventos asociados a errores técnicos.
Ma et al. (2024) (2)	n=42 participantes, EE. UU.	Ensayo clínico piloto aleatorizado	2 sesiones	Evaluar feedback basado en IA en sutura robótica	Mejora significativa en el manejo de la aguja y en el desempeño técnico durante la ejecución del procedimiento.
Yilmaz et al. (2024) (1)	n≈97 participantes	Ensayo clínico aleatorizado	Sesiones simuladas	Comparar IA vs instrucción tradicional	El uso de inteligencia artificial evidenció un desempeño igual o superior en el desarrollo de habilidades técnicas.

Yilmaz et al. (2025) (15)	n=72 participantes	Ensayo clínico crossover	Entrenamiento simulado	Evaluar combinación IA + experto	La estrategia combinada se asoció con un mayor rendimiento en el proceso de aprendizaje.
Mirchi et al. (2020) (16)	Residentes y fellows, Canadá	Experimental	Simulación VR	Evaluar tutor IA explicable	La retroalimentación automatizada contribuyó a mejorar el proceso de aprendizaje y la adquisición progresiva de habilidades.
Yilmaz et al. (2022) (17)	Participantes en simulación	Experimental	N/A	Evaluar monitoreo continuo con IA	Evaluación objetiva y continua del desempeño durante el proceso formativo.
Igaki et al. (2023) (18)	Cirujanos	Estudio de validación	N/A	Evaluar sistema automático de habilidades quirúrgicas	Alta concordancia con los parámetros establecidos por la evaluación estándar.
Kitaguchi et al. (2021) (19)	Cirujanos	Estudio de validación	N/A	Evaluar modelo 3D CNN	Nivel de precisión moderado, aproximadamente del 75 %, para la clasificación de los resultados.
Lavanchy et al. (2021) (20)	Cirujanos	Experimental	N/A	Automatizar evaluación quirúrgica	Valoración objetiva del desempeño con resultados reproducibles.
Khalid et al. (2020) (21)	n=8 cirujanos	Experimental	N/A	Evaluar acciones quirúrgicas	Detección precisa de las acciones técnicas vinculadas al desempeño quirúrgico.
Funke et al. (2019) (22)	Dataset quirúrgico	Experimental	N/A	Evaluación técnica por video	Elevado nivel de precisión para la clasificación de los resultados.
Hung et al. (2018) (23)	n=50 cirujanos	Experimental	N/A	Evaluar desempeño robótico	Correspondencia significativa con los criterios establecidos por expertos.

Chen et al. (2020) (24)	n=80 residentes	Ensayo clínico	6 meses	Evaluar habilidades laparoscópicas	Incremento significativo del rendimiento durante el proceso formativo.
Fard et al. (2018) (25)	n=40 participantes	Experimental	N/A	Evaluar habilidades robóticas	Clasificación con alto nivel de precisión.
Zia et al. (2018) (26)	n=30 participantes	Experimental	N/A	Evaluación automatizada	Alto grado de precisión en los resultados obtenidos.
Power et al. (2025) (27)	Médicos en formación	Experimental	N/A	Clasificar habilidades	Diferenciación efectiva de los elementos evaluados.
Pan et al. (2023) (28)	Dataset	Experimental	N/A	Evaluar habilidades	Nivel de precisión mayor al 85 % en los resultados obtenidos.
Ebina et al. (2022) (29)	Cirujanos	Experimental	N/A	Evaluar competencia	Identificación de métricas relevantes para la evaluación del desempeño.
Goldbraikh et al. (2022) (30)	n=25 participantes	Experimental	N/A	Evaluar sutura abierta	Proceso de evaluación automatizada factible y aplicable al contexto formativo.
Nakajima et al. (2024) (31)	Cirujanos	Experimental	N/A	Evaluar fases quirúrgicas	Utilidad de la inteligencia artificial en el análisis del desempeño.
Yanagida et al. (2025) (32)	Cirujanos	Experimental	N/A	Evaluar fases laparoscópicas	Modelo funcional para su aplicación en el contexto evaluativo.
Komatsu et al. (2024) (33)	Multicéntrico	Experimental	N/A	Evaluar cirugía gástrica	Modelo con validez para apoyar la evaluación del desempeño.
Jin et al. (2018) (34)	Cirujanos	Experimental	N/A	Detectar instrumentos	Elevado nivel de precisión en los resultados obtenidos.

Ahmidi et al. (2017) (35)	Dataset	Experimental	N/A	Reconocer gestos	Dataset validado para su uso en procesos de análisis y evaluación.
Gao et al. (2014) (36)	Dataset	Experimental	N/A	Modelar habilidades	Referencia estándar para el análisis y la evaluación del desempeño.
Yanik et al. (2023) (37)	Dataset	Experimental	N/A	Evaluación automática	Retroalimentación basada en recursos visuales para apoyar la evaluación del desempeño.
Trujillo et al. (2024) (3)	Residentes	Observacional	N/A	Evaluar IA generativa	Desempeño de la inteligencia artificial equiparable al juicio humano.
Tran et al. (2024) (38)	N/A	Observacional	N/A	Evaluar conocimientos	Nivel moderado de precisión en los resultados obtenidos.
Rezaei et al. (2023) (39)	Dataset	Experimental	N/A	Evaluar habilidades	Modelo sólido y consistente para el procesamiento de resultados.

La problemática principal de este tema se vincula con la distancia existente entre las exigencias del entrenamiento quirúrgico de posgrado y las limitaciones presentes en los entornos clínicos reales. La formación de residentes requiere sistemas de evaluación objetivos, retroalimentación permanente y espacios de práctica que favorezcan el desarrollo de habilidades técnicas sin comprometer la seguridad del paciente. No obstante, el aprendizaje en el quirófano puede estar condicionado por factores como la disponibilidad del docente, la diversidad de casos clínicos, la carga asistencial y la falta de criterios estandarizados para medir el desempeño (14).

Frente a este escenario, la inteligencia artificial surge como una herramienta con capacidad para complementar la educación quirúrgica, al facilitar mediciones automatizadas, reducir la variabilidad en la evaluación y proporcionar retroalimentación de forma más organizada. Sin embargo, sus aportes no son automáticos ni homogéneos, ya que dependen de la tecnología utilizada, del contexto de aplicación y del enfoque pedagógico que guíe su incorporación (15).

Entre los estudios con aplicación más directa en la formación quirúrgica de posgrado se encuentra el ensayo aleatorizado multicéntrico que analizó el uso de SmartCoach en colecistectomía laparoscópica.

En esta investigación, el programa de acompañamiento basado en retroalimentación asistida por inteligencia artificial se asoció con una mejora del desempeño evaluado mediante LCRF, evidenciando mejores resultados en el grupo intervenido en comparación con el grupo de autoaprendizaje al finalizar la intervención. Además, se observó un aumento importante en la obtención de la CVS, que pasó del 11 % al 78 %. La población estudiada resulta especialmente pertinente para el contexto de residencia, ya que estuvo conformada principalmente por residentes y fellows, con una distribución de 78 % y 22 %, respectivamente, en ambos grupos (14).

Estos hallazgos sugieren que la inteligencia artificial puede aportar a un componente clave del aprendizaje quirúrgico: la estandarización de criterios de seguridad y el fortalecimiento del rendimiento en procedimientos reales. Así, esta herramienta contribuye a responder a la necesidad de una formación posgradual más segura, homogénea y acorde con las demandas clínicas actuales (14).

En el ámbito de la simulación robótica, un estudio piloto aleatorizado analizó el efecto de la retroalimentación asistida por inteligencia artificial durante la realización de una tarea de sutura. La investigación incluyó participantes sin experiencia previa en cirugía robótica, quienes fueron asignados aleatoriamente a un grupo control y a un grupo de intervención que recibió retroalimentación automatizada (16). Los resultados mostraron una mejora superior en el componente relacionado con el manejo de la aguja en el grupo apoyado por IA, con una diferencia de 0,30 frente a -0,02 en el grupo control.

Este efecto fue más evidente en los participantes con menor rendimiento inicial, mientras que el grupo control recibió clips aleatorios utilizados como placebo. Desde una perspectiva educativa, este hallazgo resulta relevante, ya que sugiere que la retroalimentación automatizada podría ayudar a identificar con mayor precisión a los residentes que necesitan mayor acompañamiento durante la adquisición de habilidades técnicas (14).

En este sentido, la IA podría contribuir a una progresión más eficiente dentro de la curva de aprendizaje. Además, este estudio se relaciona con una dificultad frecuente en la formación quirúrgica: la retroalimentación en el entorno operatorio suele depender de la disponibilidad de casos, del tiempo del docente y de la carga asistencial, lo que puede ocasionar diferencias en la calidad, oportunidad y contenido de la orientación recibida. Por ello, esta evidencia se vincula directamente con las limitaciones propias de la residencia y con la necesidad de estudiar alternativas automatizadas aplicadas a la enseñanza quirúrgica (2).

En el campo de la simulación neuroquirúrgica, se identificó un ensayo clínico aleatorizado con tres grupos de comparación, en el cual se evaluó la retroalimentación en tiempo real generada por inteligencia artificial frente a la enseñanza presencial de un experto y la retroalimentación proporcionada después de la práctica. El desempeño fue valorado mediante indicadores específicos y con el instrumento OSATS.

Aunque la población estuvo compuesta principalmente por estudiantes y no por residentes médicos, los resultados aportan un fundamento teórico importante para la educación quirúrgica, al demostrar que un sistema automatizado puede orientar la práctica repetida, evaluar el rendimiento y monitorear el progreso del aprendiz.

Es por esto que, aunque la evidencia es indirecta para el contexto de posgrado, puede servir como base para el diseño de futuras intervenciones dirigidas a residentes quirúrgicos. Si estas herramientas muestran capacidad para guiar el aprendizaje en tiempo real y seguir la evolución técnica del usuario, el siguiente paso debería ser evaluarlas en programas de residencia, incluyendo desenlaces relacionados con la transferencia de habilidades, el desempeño clínico y el logro de competencias profesionales (1).

En relación con la evaluación automatizada del rendimiento dentro del quirófano, Igaki y colaboradores realizaron un estudio diagnóstico en el que emplearon un modelo de deep learning para identificar campos quirúrgicos estandarizados durante procedimientos de sigmoidectomía laparoscópica (18). A partir de este sistema se obtuvo una puntuación de confianza generada por IA, la cual mostró una correlación elevada con la escala ESSQS. Este hallazgo indica que la inteligencia artificial podría utilizarse como un recurso complementario y objetivo para valorar la calidad del desempeño técnico, e incluso como una herramienta inicial de cribado o apoyo en los procesos de evaluación (16).

Su relevancia se relaciona con una dificultad frecuente en la formación quirúrgica de posgrado: el análisis manual de videos demanda tiempo, genera costos y puede estar influido por la variabilidad entre evaluadores. Frente a esta limitación, la IA ofrece una alternativa con mayor posibilidad de aplicación a gran escala; sin embargo, su utilidad real dependerá de que los modelos conserven precisión, validez y capacidad de adaptación en contextos diferentes a aquellos donde fueron entrenados o desarrollados (17).

Otra línea de interés creciente corresponde al uso de modelos de lenguaje de gran escala para generar retroalimentación educativa. En el estudio “Surgeons vs ChatGPT”, se evidenció que, ante escenarios quirúrgicos descritos por escrito, ChatGPT identificó errores con una frecuencia comparable a la de cirujanos experimentados. Además, la retroalimentación producida fue considerada útil para apoyar la mejora de habilidades, incluyendo la valoración de residentes durante el proceso evaluativo (3).

Para el presente proyecto, estos resultados no deben interpretarse como una sustitución del docente por parte del modelo, sino como una posible herramienta de apoyo para elaborar comentarios narrativos, organizar guías de reflexión y complementar la retroalimentación formativa. Aun así, su implementación requiere supervisión profesional y mecanismos de control que permitan disminuir el riesgo de respuestas erróneas que puedan parecer correctas por su forma de presentación (3).

Al relacionar estos resultados con la problemática general del estudio, es posible identificar tres elementos centrales. En primer lugar, ya se cuenta con investigaciones de intervención que han incorporado

inteligencia artificial en procedimientos quirúrgicos reales, con una participación relevante de médicos en formación de posgrado, y que han reportado mejoras en el desempeño técnico y en indicadores vinculados con la seguridad.

En segundo lugar, la IA parece contribuir de manera importante a la generación de retroalimentación objetiva, individualizada y basada en datos, especialmente en actividades técnicas, con un posible beneficio mayor en aquellos residentes que presentan un rendimiento inicial más bajo (18).

En tercer lugar, el desafío actual no se limita a demostrar que los modelos de IA pueden clasificar o estimar habilidades, ya que diversos estudios han mostrado esta capacidad. El verdadero reto consiste en determinar si su incorporación en currículos basados en competencias puede lograr una transferencia sostenida hacia la práctica clínica, fortalecer la seguridad del paciente y disminuir la carga docente, sin comprometer el desarrollo del razonamiento clínico ni del juicio profesional.

4. RESULTADOS

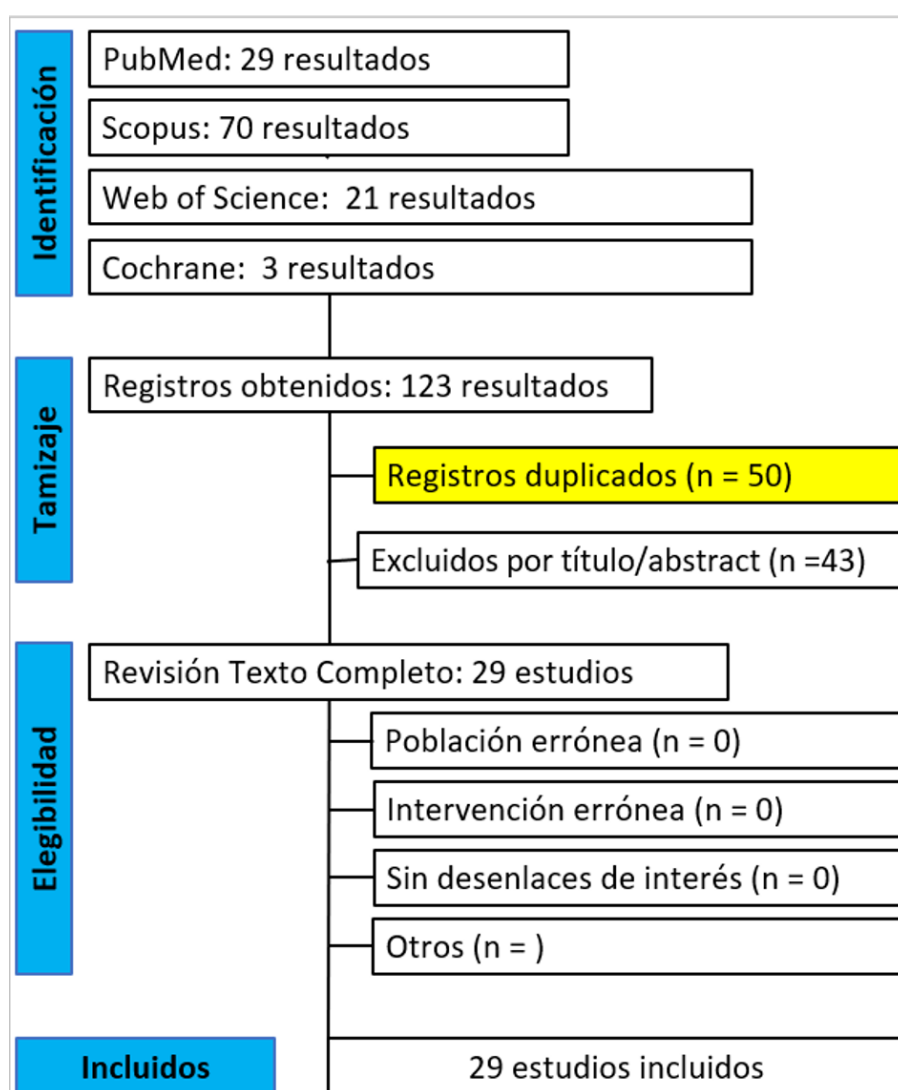


Figura 1. Diagrama de flujo PRISMA de tamizaje y selección

Figura 1. Diagrama de flujo PRISMA de tamizaje y selección.

Los estudios incluidos en esta revisión muestran una tendencia común: la inteligencia artificial puede aportar de manera significativa al fortalecimiento de varios componentes del aprendizaje quirúrgico en el posgrado. En cuanto al desarrollo de competencias técnicas, diferentes investigaciones experimentales han reportado que los residentes que entrenan con simuladores asistidos por IA alcanzan resultados similares o incluso mejores que aquellos obtenidos con métodos tradicionales. Asimismo, estas herramientas se han vinculado con una reducción de errores durante la práctica y con una mayor eficiencia en la ejecución de procedimientos específicos, evidenciada en una disminución del tiempo necesario para completar determinadas tareas quirúrgicas.

También se ha observado que las tecnologías basadas en inteligencia artificial pueden favorecer el refinamiento de la técnica operatoria y mejorar la precisión de los movimientos. Este efecto se relaciona especialmente con la retroalimentación automatizada y con la posibilidad de evaluar las habilidades de forma más objetiva y sistemática (Hung et al., Mirchi et al.) (16, 23). Además, los modelos predictivos basados en machine learning han demostrado utilidad para identificar patrones de rendimiento y anticipar posibles desenlaces clínicos. Esto permite reconocer tempranamente dificultades formativas y planificar intervenciones educativas más oportunas en residentes o médicos en proceso de especialización (Loftus et al., Senders et al.) (40, 41).

En cuanto a la evaluación del desempeño, la inteligencia artificial permite construir mediciones más objetivas, estables y comparables entre distintos aprendices. Esta capacidad resulta importante porque ayuda a reducir la variabilidad que suele presentarse cuando la valoración depende únicamente del criterio del evaluador. De este modo, su incorporación podría mejorar la confiabilidad de los procesos de medición de competencias quirúrgicas y contribuir a sistemas de certificación más homogéneos, claros y sustentados en evidencia. Además, la retroalimentación automatizada ofrece al residente información inmediata sobre su ejecución, lo que facilita el reconocimiento de errores, la corrección temprana de fallos técnicos y un avance más ordenado dentro del proceso de aprendizaje.

La evidencia revisada también resalta que estas herramientas pueden incrementar la eficiencia del entrenamiento quirúrgico. Varios estudios señalan que los recursos basados en inteligencia artificial permiten aprovechar mejor el tiempo destinado a la práctica, ya que disminuyen parcialmente la dependencia de la supervisión directa y favorecen ejercicios más repetidos, específicos y adaptados a las necesidades individuales del residente. Este aporte resulta especialmente valioso en contextos donde el tiempo de formación es limitado y la disponibilidad de docentes, tutores o especialistas para acompañar el aprendizaje no siempre es suficiente.

5. DISCUSIÓN

La evidencia recopilada en esta revisión sistemática permite afirmar que la inteligencia artificial tiene un amplio potencial para transformar diversos aspectos de la formación quirúrgica de posgrado. Sus

contribuciones más importantes se relacionan con la evaluación más objetiva del desempeño, la adaptación del aprendizaje a las características individuales del residente y el uso más eficiente del tiempo destinado al entrenamiento. Sin embargo, estos resultados deben interpretarse con cautela, debido a que aún existen limitaciones metodológicas, variabilidad entre los estudios incluidos y desafíos asociados a la aplicación de estas tecnologías en escenarios educativos y clínicos reales.

Uno de los aportes más destacados de la inteligencia artificial es su capacidad para apoyar evaluaciones basadas en datos cuantificables, reproducibles y comparables. Esto marca una diferencia frente a los métodos tradicionales, que con frecuencia dependen del juicio subjetivo del evaluador. Gracias a esta característica, la IA podría mejorar la precisión con la que se valoran las competencias quirúrgicas y facilitar la identificación de patrones de desempeño que no siempre son evidentes mediante la observación directa. Desde este enfoque, su uso se alinea con los modelos de formación basada en competencias, en los cuales el avance del residente requiere seguimiento continuo, criterios definidos y mediciones más uniformes.

Asimismo, la posibilidad de personalizar el aprendizaje representa una ventaja relevante de estas herramientas. Los sistemas apoyados en inteligencia artificial pueden modificar contenidos, niveles de complejidad y actividades prácticas de acuerdo con el rendimiento de cada aprendiz. Esto permite emplear de forma más adecuada el tiempo de formación y favorecer un aprendizaje más significativo, ajustado a las necesidades reales del residente. Esta característica resulta especialmente importante en la educación quirúrgica, porque no todos los médicos en formación progresan al mismo ritmo ni requieren el mismo tipo de apoyo durante el desarrollo de sus habilidades.

La evidencia analizada en esta revisión sistemática muestra que la inteligencia artificial puede convertirse en un recurso relevante para fortalecer la formación quirúrgica de posgrado. Su aporte principal se relaciona con la posibilidad de mejorar la objetividad de la evaluación, ajustar el aprendizaje al desempeño individual del residente y optimizar el tiempo destinado al entrenamiento. Sin embargo, estos resultados no deben asumirse de manera absoluta, ya que la evidencia aún presenta limitaciones importantes. Entre ellas destacan las diferencias metodológicas entre los estudios, la variedad de intervenciones evaluadas y las dificultades para trasladar estas herramientas a contextos clínicos y educativos reales.

Uno de los beneficios más importantes de la inteligencia artificial es su capacidad para generar valoraciones basadas en información cuantificable, comparable y reproducible. Esto representa una ventaja frente a los métodos tradicionales de evaluación, los cuales suelen estar influenciados por la percepción y experiencia del evaluador. A través de estos sistemas, es posible valorar con mayor precisión las competencias quirúrgicas e identificar patrones de desempeño que podrían no ser detectados mediante la observación directa. En este sentido, la IA se articula con los enfoques de educación basada en

competencias, donde el avance del residente requiere seguimiento permanente, criterios claros y mediciones objetivas.

La personalización del aprendizaje también constituye un aspecto relevante dentro de estas aplicaciones. Las herramientas basadas en inteligencia artificial pueden adaptar las actividades, los contenidos y el nivel de complejidad del entrenamiento según las necesidades y el rendimiento de cada residente. Esta característica permite aprovechar mejor los espacios formativos y promover un aprendizaje más ajustado a las dificultades reales del médico en formación. En cirugía, este punto adquiere especial importancia, debido a que los residentes no desarrollan sus habilidades técnicas y cognitivas al mismo ritmo, ni requieren el mismo tipo de acompañamiento durante todo el proceso educativo.

A pesar de estos aportes, la evidencia disponible presenta debilidades que deben ser tomadas en cuenta. Una de las principales es la heterogeneidad de los estudios incluidos, tanto por sus diseños metodológicos como por las tecnologías utilizadas y los desenlaces seleccionados para medir los resultados. Esta diversidad dificulta la comparación directa entre investigaciones y limita la posibilidad de establecer conclusiones cuantitativas firmes. Además, muchos estudios trabajan con muestras reducidas y se realizan en ambientes simulados, lo que disminuye la certeza sobre la aplicabilidad de sus hallazgos en escenarios quirúrgicos reales.

También es necesario considerar que muchos modelos de inteligencia artificial cuentan con escasa validación externa. En varios estudios, los algoritmos se diseñan, entrenan y prueban dentro de un mismo entorno institucional o con grupos muy específicos de participantes. Esto limita la seguridad sobre su desempeño en otros programas de residencia, especialidades quirúrgicas o contextos asistenciales. Por esta razón, todavía existen dudas sobre su capacidad de generalización y sobre la posibilidad de implementarlos de forma amplia y sostenible en la educación quirúrgica de posgrado.

Desde una perspectiva ética, la incorporación de inteligencia artificial en la educación médica requiere un análisis cuidadoso, ya que su uso no está libre de riesgos. La dependencia excesiva de sistemas automatizados podría limitar el desarrollo del razonamiento clínico, la autonomía profesional y la capacidad de tomar decisiones fundamentadas durante la formación. Además, algunos algoritmos funcionan con baja transparencia, lo que dificulta comprender cómo generan sus resultados y puede afectar la confianza de docentes y residentes. A ello se suma la posibilidad de que estas tecnologías profundicen brechas educativas, especialmente entre instituciones con mayor disponibilidad de recursos y aquellas que enfrentan limitaciones económicas, tecnológicas o de infraestructura.

Por esta razón, la inteligencia artificial debe entenderse como una herramienta complementaria dentro del proceso formativo, y no como un sustituto de la enseñanza tradicional ni de la supervisión docente. La guía del profesor, el acompañamiento clínico, el criterio profesional y la interacción humana siguen siendo componentes esenciales en la formación quirúrgica. En consecuencia, su implementación debe realizarse

de manera crítica, gradual y equilibrada, integrada a un modelo pedagógico sólido que aproveche las ventajas tecnológicas sin desplazar el papel central del docente en el aprendizaje médico.

6. APLICACIONES PRÁCTICAS O FUTURAS LÍNEAS DE INVESTIGACIÓN

La presente investigación bibliográfica tiene como aplicación práctica servir de base para la incorporación progresiva de herramientas de inteligencia artificial en los programas de educación quirúrgica de posgrado. Su uso puede fortalecer el entrenamiento mediante simuladores quirúrgicos inteligentes, sistemas de retroalimentación automatizada y plataformas de aprendizaje adaptativo que permitan evaluar de forma objetiva el desempeño técnico de los médicos en formación.

Además, esta revisión puede orientar a instituciones académicas y hospitalarias en el diseño de estrategias educativas más personalizadas, seguras y estandarizadas, complementando la enseñanza tradicional con tecnologías que ayuden a identificar errores, mejorar habilidades quirúrgicas y optimizar las curvas de aprendizaje.

Como futuras líneas de investigación, se recomienda desarrollar estudios con muestras más amplias, diseños multicéntricos y seguimiento a largo plazo, que permitan evaluar la efectividad real de la inteligencia artificial en diferentes especialidades quirúrgicas. También sería importante investigar sus implicaciones éticas, la aceptación por parte de docentes y residentes, los costos de implementación y su impacto directo en la seguridad del paciente.

7. CONCLUSIONES

La inteligencia artificial se perfila como una herramienta de apoyo importante para la formación quirúrgica de posgrado, sobre todo porque permite incorporar procesos de enseñanza, práctica y evaluación basados en datos más objetivos. A partir de la evidencia revisada, se observa que su uso en simuladores, sistemas de análisis del desempeño y plataformas de aprendizaje puede contribuir al desarrollo de habilidades técnicas, a la disminución de errores y a una mejor organización del proceso de entrenamiento de los residentes.

Uno de los aportes más relevantes de esta revisión es que muestra cómo la inteligencia artificial puede ayudar a superar algunas limitaciones frecuentes de la educación quirúrgica tradicional. Entre ellas se encuentran la evaluación subjetiva del rendimiento, la variabilidad en las oportunidades de práctica y la dificultad para ofrecer retroalimentación individualizada de manera constante. En este sentido, los hallazgos pueden ser útiles para docentes, investigadores y responsables de programas de residencia que buscan mejorar la calidad del entrenamiento quirúrgico.

Para otros campos de la educación médica, estos resultados también tienen valor, ya que demuestran que las tecnologías basadas en inteligencia artificial no solo pueden aplicarse en cirugía, sino también en otras áreas donde sea necesario evaluar competencias prácticas, identificar errores y adaptar el aprendizaje

a las necesidades de cada estudiante. De forma indirecta, estos avances también pueden beneficiar al público en general, al favorecer la formación de profesionales mejor preparados y con mayor capacidad para realizar procedimientos de manera segura.

Los resultados encontrados amplían lo descrito en estudios previos, debido a que no se limitan únicamente al uso de simuladores, sino que también destacan la utilidad de la retroalimentación automatizada, la evaluación objetiva y la personalización del aprendizaje. Esto refuerza la idea de que la inteligencia artificial puede complementar la enseñanza tradicional, siempre que se mantenga la supervisión docente, el juicio clínico y la experiencia práctica como elementos centrales de la formación quirúrgica.

Finalmente, aunque la evidencia disponible es alentadora, todavía debe interpretarse con prudencia. La heterogeneidad de los estudios, los tamaños muestrales reducidos y la limitada validación externa de algunas herramientas indican que aún se necesitan investigaciones más amplias. Por ello, sería conveniente desarrollar estudios multicéntricos, con seguimiento a largo plazo y criterios de evaluación estandarizados, que permitan conocer con mayor precisión el verdadero impacto de la inteligencia artificial en la formación quirúrgica, la seguridad del paciente y la calidad de la atención médica.

8. AGRADECIMIENTOS

Dedico un especial agradecimiento a mis hijos, Oliver y Matías, por su comprensión, paciencia y constante motivación durante el desarrollo de este trabajo de posgrado. Su apoyo y presencia han sido una fuente permanente de inspiración para alcanzar esta meta académica (Moreta E, 2026).

Expreso mi más profundo agradecimiento a mi madre, por estar siempre a mi lado y brindarme su apoyo incondicional; y a mis hijos, quienes constituyen mi mayor inspiración, mi motor y la fuerza que necesito para seguir adelante. A ellos dedico este logro, con infinito amor y gratitud (Pazmiño Y, 2026).

9. FINANCIACIÓN

La presente investigación no recibió financiamiento de instituciones públicas, privadas, organizaciones gubernamentales, universidades, centros de investigación, becas ni proyectos aprobados. El estudio fue desarrollado con recursos propios de los autores.

10. CONFLICTO DE INTERESES

Los autores declaran no tener ningún conflicto de intereses relacionado con el desarrollo, los resultados o la publicación de la presente investigación.

11. CONTRIBUCIÓN DE AUTORÍA

	Autor 1.	Autor 2
Participar activamente en:		
conceptualización	X	
Análisis formal		x
Adquisición de fondos	X	
Investigación		x
Metodología	X	
Administración del proyecto		x
Recursos	X	
Redacción –borrador original		X
Redacción –revisión y edición	X	
La discusión de los resultados	X	x
Revisión y aprobación de la versión final del trabajo.	X	x

REFERENCIAS BIBLIOGRÁFICAS

1. Yilmaz R, Bakhaidar M, Alsayegh A, Abou Hamdan N, Fazlollahi AM, Tee T, et al. Real-Time multifaceted artificial intelligence vs In-Person instruction in teaching surgical technical skills: a randomized controlled trial. Sci Rep. 2024.
URL: <https://pmc.ncbi.nlm.nih.gov/articles/PMC11219907/>

2. Ma R, Kiyasseh D, Laca JA, et al. Artificial Intelligence-Based Video Feedback to Improve Novice Performance on Robotic Suturing Skills: A Pilot Study. J Endourol. 2024;38(8):884–891.

DOI: <https://doi.org/10.1089/end.2023.0328>.

URL: <https://pmc.ncbi.nlm.nih.gov/articles/PMC11947633/>

3. Trujillo CJ, Vela Ulloa J, Escalona Vivas G, et al. Surgeons vs ChatGPT: Assessment and Feedback Performance Based on Real Surgical Scenarios. *J Surg Educ.* 2024;81(7):960–966. DOI: <https://doi.org/10.1016/j.jsurg.2024.03.012>. URL: <https://pubmed.ncbi.nlm.nih.gov/38749814/>
4. Chan KS, Zary N. Applications of artificial intelligence in surgical simulation. *Surg Innov.* 2022;29(3):321–329. Disponible en: <https://pubmed.ncbi.nlm.nih.gov/34962116/>
5. Frank JR, Mungroo R, Ahmad Y, et al. Toward a definition of competency-based education in medicine: a systematic review of published definitions. *Med Teach.* 2010;32(8):631–637. URL: <https://pubmed.ncbi.nlm.nih.gov/20662573/>
6. Hattie J, Timperley H. The Power of Feedback. *Rev Educ Res.* 2007;77(1):81–112. DOI: <https://doi.org/10.3102/003465430298487>. URL: <https://journals.sagepub.com/doi/10.3102/003465430298487>
7. Martin JA, Regehr G, Reznick R, et al. Objective structured assessment of technical skill (OSATS) for surgical residents. *Br J Surg.* 1997;84(2):273–278. URL: <https://pubmed.ncbi.nlm.nih.gov/9052454/>
8. Ahmidi N, Tao L, Sefati S, et al. A dataset and benchmarks for segmentation and recognition of gestures in robotic surgery. *IEEE Trans Biomed Eng.* 2017;64(9):2025–2041. DOI: <https://doi.org/10.1109/TBME.2016.2647680>. URL: <https://pmc.ncbi.nlm.nih.gov/articles/PMC5559351/>
9. Lefor AK, Harada K, Dosis A, Mitsuishi M. Motion analysis of the JHU-ISI Gesture and Skill Assessment Working Set using Robotics Video and Motion Assessment Software. *Int J Comput Assist Radiol Surg.* 2020 Dec;15(12):2017–2025. DOI: 10.1007/s11548-020-02259-z.
10. Gao Y, Vedula SS, Reiley CE, et al. JHU-ISI Gesture and Skill Assessment Working Set (JIGSAWS): A surgical activity dataset for human motion modeling. *MICCAI Workshop M2CAI.* 2014. URL: <https://cirl.lcsr.jhu.edu/wp-content/uploads/2015/11/JIGSAWS.pdf>
11. Ward TM, Mascagni P, Ban Y, et al. Computer vision in surgery. *J Surg Educ.* 2021;78(2):325–332. Disponible en: <https://pubmed.ncbi.nlm.nih.gov/33144155/>
12. Goldenberg MG. Surgical artificial intelligence in urology: educational applications. *Urol Clin North Am.* 2024;51(1):105–115. Disponible en: <https://doi.org/10.1016/j.ucl.2023.06.003>
13. Kirubarajan A, Young D, Khan S, Crasto N, Sobel M, Sussman D. Artificial intelligence and surgical education: a systematic scoping review of interventions. *J Surg Educ.* 2022;79(2):500–515. Disponible en: <https://pubmed.ncbi.nlm.nih.gov/34756807/>
14. Wu S, Tang M, Wang X, et al. Impact of an AI-based laparoscopic cholecystectomy coaching program on the surgical performance: a randomized controlled trial. *Int J Surg.* 2024. DOI: <https://doi.org/10.1097/JS9.0000000000001798>. URL: <https://pmc.ncbi.nlm.nih.gov/articles/PMC11634122/>

15. Yilmaz R., Alsayegh A., Bakhaidar M. *et al.* Combining real-time AI and in-person expert instruction in simulated surgical skills training - Randomized crossover trial. *npj Artif. Intell.* **1**, 36 (2025). <https://doi.org/10.1038/s44387-025-00032-8>
16. Mirchi N, Bissonnette V, Yilmaz R, Ledwos N, Winkler-Schwartz A, Del Maestro RF. The Virtual Operative Assistant: An explainable artificial intelligence tool for simulation-based training in surgery and medicine. *PLoS One*. 2020;15(2):e0229596.
Disponibile en: <https://doi.org/10.1371/journal.pone.0229596>
17. Yilmaz R, Winkler-Schwartz A, Mirchi N, Reich A, Christie S, Tran DH, et al. Continuous monitoring of surgical bimanual expertise using deep neural networks in virtual reality simulation. *npj Digit Med*. 2022;5(1):54.
Disponibile en: <https://doi.org/10.1038/s41746-022-00596-8>
18. Igaki T, et al. Automatic Surgical Skill Assessment System Based on Deep Learning. *JAMA Surg*. 2023. URL: <https://pmc.ncbi.nlm.nih.gov/articles/PMC10248810/>
19. Kitaguchi D, Takeshita N, Matsuzaki H, Igaki T, Hasegawa H, Ito M. Development and Validation of a 3-Dimensional Convolutional Neural Network for Automatic Surgical Skill Assessment Based on Spatiotemporal Video Analysis. *JAMA Netw Open*. 2021;4(8):e2120786.
doi:10.1001/jamanetworkopen.2021.20786
20. Lavanchy J L., Zindel J., Kirtac K. *et al.* Automation of surgical skill assessment using a three-stage machine learning algorithm. *Sci Rep* **11**, 5197 (2021). <https://doi.org/10.1038/s41598-021-84295-6>
21. Khalid S, Goldenberg M, Grantcharov T, Taati B, Rudzicz F. Evaluation of Deep Learning Models for Identifying Surgical Actions and Measuring Performance. *JAMA Netw Open*. 2020 Mar 2;3(3):e201664. doi: 10.1001/jamanetworkopen.2020.1664. PMID: 32227178; PMCID: PMC12124734.
22. Funke I, Mees ST, Weitz J, Speidel S. Video-based surgical skill assessment using 3D convolutional neural networks. *Int J Comput Assist Radiol Surg*. 2019 Jul;14(7):1217-1225. doi: 10.1007/s11548-019-01995-1. Epub 2019 May 18. PMID: 31104257.
23. Hung AJ, Chen J, Che Z, Nilanon T, Jarc A, Titus M, et al. Utilizing machine learning and automated performance metrics to evaluate robot-assisted radical prostatectomy performance and predict outcomes. *J Endourol*. 2018;32(5):438–444. DOI: <https://doi.org/10.1089/end.2018.0035>
24. Chen J, Cheng N, Cacciamani G, Oh P, Lin-Brandt M, Remulla D, Gill IS, Hung AJ. Objective assessment of robotic surgical technical skill: a systematic review. *J Urol*. 2019;201(3):461–469. Disponibile en: <https://pubmed.ncbi.nlm.nih.gov/30053510/>
25. Fard, M. J., Ameri, S., Darin Ellis, R., Chinnam, R. B., Pandya, A. K., & Klein, M. D. (2018). Automated robot-assisted surgical skill evaluation: Predictive analytics approach. *The international journal of medical robotics + computer assisted surgery : MRCAS*, 14(1), 10.1002/rcs.1850. <https://doi.org/10.1002/rcs.1850>
26. Zia, A., & Essa, I. (2018). Automated surgical skill assessment in RMIS training. *International journal of computer assisted radiology and surgery*, 13(5), 731–739. <https://doi.org/10.1007/s11548-018-1735-5>
27. Power, D., Burke, C., Madden, M.G. et al. Automated assessment of simulated laparoscopic surgical skill performance using deep learning. *Sci Rep* **15**, 13591 (2025). <https://doi.org/10.1038/s41598-025-96336-5>

28. Pan, M., Wang, S., Li, J., Li, J., Yang, X., & Liang, K. (2023). An Automated Skill Assessment Framework Based on Visual Motion Signals and a Deep Neural Network in Robot-Assisted Minimally Invasive Surgery. *Sensors*, 23(9), 4496. <https://doi.org/10.3390/s23094496>
29. Ebina, K., Abe, T., Hotta, K., Higuchi, M., Furumido, J., Iwahara, N., Kon, M., Miyaji, K., Shibuya, S., Lingbo, Y., Komizunai, S., Kurashima, Y., Kikuchi, H., Matsumoto, R., Osawa, T., Murai, S., Tsujita, T., Sase, K., Chen, X., Konno, A., Shinohara, N. (2022). Automatic assessment of laparoscopic surgical skill competence based on motion metrics. *PloS one*, 17(11), e0277105. <https://doi.org/10.1371/journal.pone.0277105>
30. Goldbraikh, A., D'Angelo, A. L., Pugh, C. M., & Laufer, S. (2022). Video-based fully automatic assessment of open surgery suturing skills. *International journal of computer assisted radiology and surgery*, 17(3), 437–448. <https://doi.org/10.1007/s11548-022-02559-6>
31. Nakajima, K., Kitaguchi, D., Takenaka, S. *et al.* Automated surgical skill assessment in colorectal surgery using a deep learning-based surgical phase recognition model. *Surg Endosc* **38**, 6347–6355 (2024). <https://doi.org/10.1007/s00464-024-11208-9>
32. Yanagida, Y., Takenaka, S., Kitaguchi, D., Hamano, S., Tanaka, A., Mitarai, H., Suzuki, R., Sasaki, K., Takeshita, N., Ishimaru, T., Fujishiro, J., & Ito, M. (2025). Surgical skill assessment using an AI-based surgical phase recognition model for laparoscopic cholecystectomy. *Surgical endoscopy*, 39(8), 5018–5026. <https://doi.org/10.1007/s00464-025-11903-1>
33. Komatsu, M., Kitaguchi, D., Yura, M., Takeshita, N., Yoshida, M., Yamaguchi, M., Kondo, H., Kinoshita, T., & Ito, M. (2024). Automatic surgical phase recognition-based skill assessment in laparoscopic distal gastrectomy using multicenter videos. *Gastric cancer : official journal of the International Gastric Cancer Association and the Japanese Gastric Cancer Association*, 27(1), 187–196. <https://doi.org/10.1007/s10120-023-01450-w>
34. Jin, A., Yeung, S., Jopling, J., Krause, J., Azagury, D., Milstein, A., & Fei-Fei, L. (2018). *Tool detection and operative skill assessment in surgical videos using region-based convolutional neural networks*. arXiv. <https://doi.org/10.48550/arXiv.1802.08774>
35. Ahmidi, N., Tao, L., Sefati, S., Gao, Y., Lea, C., Haro, B. B., Zappella, L., Khudanpur, S., Vidal, R., & Hager, G. D. (2017). A Dataset and Benchmarks for Segmentation and Recognition of Gestures in Robotic Surgery. *IEEE transactions on bio-medical engineering*, 64(9), 2025–2041. <https://doi.org/10.1109/TBME.2016.2647680>
36. Gao, Y., Vedula, S. S., Reiley, C. E., Ahmidi, N., Varadarajan, B., Lin, H. C., Tao, L., Zappella, L., Béjar, B., Yuh, D. D., Chen, C. C. G., Vidal, R., Khudanpur, S., & Hager, G. D. (2014). *The JHU-ISI gesture and skill assessment working set (JIGSAWS): A surgical activity dataset for human motion modeling*. In *Modeling and Monitoring of Computer Assisted Interventions (M2CAI) – MICCAI Workshop*. <https://cirl.lcsr.jhu.edu/wp-content/uploads/2015/11/JIGSAWS.pdf>
37. Yanik, E., Kruger, U., Intes, X. *et al.* Video-based formative and summative assessment of surgical tasks using deep learning. *Sci Rep* **13**, 1038 (2023). <https://doi.org/10.1038/s41598-022-26367-9>
38. Tran, C. G., Chang, J., Sherman, S. K., & De Andrade, J. P. (2024). Performance of ChatGPT on American Board of Surgery In-Training Examination Preparation Questions. *The Journal of surgical research*, 299, 329–335. <https://doi.org/10.1016/j.jss.2024.04.060>
39. Rezaei, A., Ahmadi, M. J., Molaei, A., & Hamid, H. (2023). *Video-based surgical skill assessment using tree-based Gaussian process classifier*. arXiv. <https://doi.org/10.48550/arXiv.2312.10208>

40. Loftus TJ, Tighe PJ, Filiberto AC, et al. Artificial intelligence and surgical decision-making. *Ann Surg.* 2020;272(1):70–80.
Disponibile en: <https://pubmed.ncbi.nlm.nih.gov/32097206/>
41. Senders JT, Staples PC, Karhade AV, et al. Machine learning in neurosurgery. *Neurosurgery.* 2018;83(2):181–192.
Disponibile en: <https://pubmed.ncbi.nlm.nih.gov/29506006/>