



TRABAJO DE FIN DE GRADO

Grado en Medicina



Universidad de Valladolid

DEBILIDADES Y FORTALEZAS DEL USO DE INTELIGENCIA ARTIFICIAL PARA EL DIAGNÓSTICO PRECOZ DE ENFERMEDADES RARAS

UNA REVISIÓN BIBLIOGRÁFICA

ANA BESTARD CASSINELLO

Tutores:

Diego Sánchez Romero
Sergio Díez Hermano

Índice

Resumen	2
Palabras clave	2
Abstract.....	2
Keywords	3
Glosario de siglas y abreviaturas	3
Introducción.....	3
Enfermedades raras.....	3
Inteligencia Artificial en Medicina	4
Justificación y objetivos	6
Justificación	6
Objetivo General.....	7
Objetivos específicos	7
Materiales y métodos	7
Estrategia de búsqueda.....	7
Criterios de selección.....	8
Evaluación de la calidad.....	9
Resultados	9
Cómo afecta la limitación de la disponibilidad de datos al entrenamiento de la IA.....	9
Análisis de aspectos del diagnóstico precoz de enfermedades raras que podrían beneficiarse del uso de IAs.....	10
Consideraciones éticas que plantea la llegada de la IA para el ejercicio de la profesión médica.....	12
Discusión	13
Direcciones futuras.....	14
Conclusión	15
Bibliografía.....	16
Anexos.....	17

Resumen

Las enfermedades raras suponen más de 400 millones de casos en el mundo. El interés en mejorar su diagnóstico no ha dejado de aumentar desde los años 90, momento en que empezaron a ganar peso para la salud pública. Actualmente, la Inteligencia Artificial (IA) ofrece una alternativa viable para conseguir un salto cualitativo en los diagnósticos, gracias al análisis de datos masivos guiado por el aprendizaje automático.

El objetivo general de este trabajo es analizar la posible aplicación de la IA en este campo y la limitación que supone la reducida disponibilidad de datos, haciendo énfasis en los beneficios potenciales del diagnóstico precoz y las consideraciones éticas que supone la llegada de la IA a la ciencia médica.

Tras realizar una búsqueda bibliográfica en PubMed, Cochrane y Ovid, la selección final consistió en 27 artículos. Se priorizó que fueran recientes, de 2020 en adelante, y de primer cuartil. También se revisó la Legislación Europea sobre IA.

Esta revisión ha permitido comprobar que IA puede mejorar el diagnóstico de enfermedades raras, reduciendo tiempos y aumentando la precisión, aunque enfrenta desafíos relacionados con la disponibilidad y calidad de datos, así como sesgos de selección, problemas de interpretabilidad y necesidad de regulación.

Palabras clave

Inteligencia Artificial, Aprendizaje Profundo, Aprendizaje Automático, Enfermedades Raras, Ética.

Abstract

Rare diseases affect over 400 million people worldwide. Interest in improving their diagnosis has steadily increased since the 1990s, when they began to gain prominence in public health. Currently, Artificial Intelligence (AI), offers a viable alternative for achieving a qualitative leap in diagnostics, thanks to the analysis of large-scale data guided by machine learning.

The main goal of this work is to examine the potential application of AI in this field and the obstacles posed by the limited availability of data, with an emphasis on the benefits of early diagnosis, and the ethical considerations arising from the integration of the AI into medical science.

PubMed, Cochrane and Ovid databases were searched for relevant literature. The final selection consisted of 27 articles. Priority was given to those that were recent (from 2020 onwards), and from the first quartile. European Legislation on AI was also reviewed.

This review has shown that AI can improve the diagnosis of rare diseases by reducing diagnostic time and increasing accuracy. However, it still faces challenges related to data availability and quality, biases, interpretability issues, and the need for regulation.

Keywords

Artificial Intelligence, Deep Learning, Machine Learning, Rare Diseases, Ethics.

Glosario de siglas y abreviaturas

IA (Inteligencia Artificial), ML (Machine Learning / Aprendizaje Automático), DL (Deep Learning / Aprendizaje Profundo), OE (Objetivo Específico), HPO (Human Phenotype Ontology).

Introducción

Enfermedades raras

Una enfermedad se considera rara en Europa si la padecen menos de 5 de cada 10.000 personas (1). Aunque de manera individual representan un colectivo muy pequeño, y su prevalencia exacta se desconoce, se estima que globalmente más de 400 millones de personas se ven afectadas por este tipo de enfermedades, según la Organización Mundial de la Salud (2)(3).

Los términos enfermedades raras y enfermedades huérfanas son sinónimos. Hasta el año 2019, el número de medicamentos huérfanos aprobados por la Administración de Alimentos y Medicamentos de EE.UU. (FDA) ascendía a 782 (4).

A partir de la década de los 90, las enfermedades raras han ido ocupando un lugar destacado en los programas de salud y en la opinión pública, con el consiguiente aumento de publicaciones científicas (5). Dependiendo de la clasificación, se calcula que hay aproximadamente 7000 enfermedades raras, y se prevé que a lo largo de los próximos años se descubrirán nuevos tipos (6).

Las enfermedades raras son un problema de salud pública. Acertar con el diagnóstico es una limitación significativa para el posterior tratamiento de las enfermedades raras. Los pacientes afectados normalmente pasan años recibiendo diagnósticos equivocados, con un retraso medio de 8 años, hasta que son correctamente diagnosticados (7). Alrededor del 80% de las enfermedades raras son de origen genético y se manifiestan predominantemente durante la edad pediátrica, lo que refuerza la importancia de un cribado genético precoz (8).

Inteligencia Artificial en Medicina

La Inteligencia Artificial (IA) se centra en llevar a cabo, de forma computacional, operaciones que habitualmente necesitarían de la capacidad cognitiva humana; como el aprendizaje, la resolución de problemas y la toma de decisiones. El aprendizaje automático (o *machine learning*, ML) es la capacidad de los ordenadores de aprender de la experiencia y mejorar su desempeño, sin ser programadas de forma explícita para ello.

El aprendizaje automático se basa en el desarrollo de algoritmos y modelos que son capaces de reconocer patrones y elaborar predicciones a partir de datos sin seguir instrucciones específicas. A medida que mejora su entrenamiento, se obtienen resultados cada vez más acertados, pudiendo así resolver cuestiones complejas, lo que aumenta el rendimiento y la capacidad de adaptación (9).

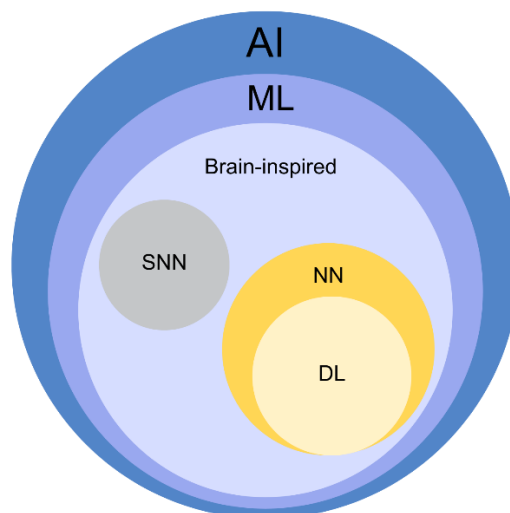


Figura 1: Correlación entre inteligencia artificial (AI), machine learning (ML), redes neuronales (NN), redes neuronales espigadas (SNN) y deep learning (DL)
Reproducido de(10).

La IA se basa en el entrenamiento de dichos algoritmos de aprendizaje automático (Figura 1). En particular, su desarrollo en los últimos años ha venido motivado por los avances en redes neurales, lo que se conoce como *deep learning* o aprendizaje profundo, debido a que son algoritmos que pueden comprender directamente datos no estructurados, como imágenes, vídeos o lenguaje hablado (10).

Existen tres enfoques principales a la hora de entrenar algoritmos de ML (Figura 2) (11):

- Aprendizaje supervisado: requiere datos etiquetados para el proceso de aprendizaje. El algoritmo produce un modelo predictivo para predicción numérica (regresión lineal, *K-Nearest Neighbors*, *gradient boosting*), clasificación (regresión logística, *random forest*, máquina de soporte vectorial, árbol de decisión) o ambos (red neuronal).
- Aprendizaje no supervisado: no necesita datos etiquetados. Produce un modelo descriptivo. Se emplea para realizar asociaciones fenómicas/genómicas a gran escala (algoritmos de aprendizaje de reglas de asociación), agrupamiento (*K-means*, modelo de espacio vectorial), o para aprender representaciones de dimensionalidad reducida (red neuronal *autoencoder*).
- Aprendizaje por refuerzo: aprende continuamente de la interacción con un entorno, para optimizar la respuesta y minimizar el riesgo de error (*Q-learning*, modelo de decisión de Markov). Estas decisiones se fundamentan en dos aspectos: resultado retrasado y búsqueda por prueba y error para ajustar sus futuras acciones.

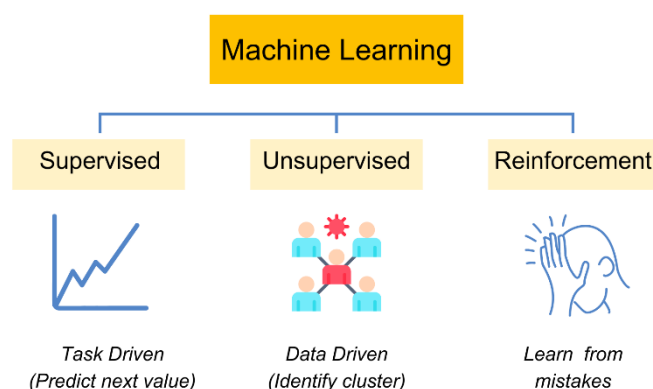


Figura 2: Enfoques del machine learning. Reproducido de (11)

La posibilidad de analizar cantidades masivas de información y de tomar decisiones con agilidad y precisión abre un amplio abanico de oportunidades relativas al cuidado de la

salud: desde la predicción y prevención de enfermedades hasta la organización intrahospitalaria, pasando por el procesamiento de imágenes médicas (12).

El aprendizaje supervisado a través del empleo de redes neuronales es, actualmente, la rama en la que más se está centrando la investigación de posibles aplicaciones de la IA para la profesión médica. Este sistema ya resulta útil en tareas de reconocimiento de imágenes, por ejemplo, para realizar estimaciones sobre el riesgo de cáncer de pulmón, y optimizar su detección, programando exploraciones de tomografía computarizada. También se utilizan para detectar irregularidades en las endoscopias. Otras funciones de la IA son la detección de resistencia a antibióticos y el descubrimiento de nuevos fármacos (13).

La precisión con la que interpreta imágenes radiológicas de diferentes áreas y el uso de modelos están ayudando a mejorar los diagnósticos. También es posible analizar imágenes médicas de manera más eficiente mediante algoritmos de aprendizaje profundo. Actualmente, se está comenzando a aplicar sistemas como AlexNet, ResNet50, y MobileNetv2 a la Dermatología y en el campo de la telemedicina para detectar características clínicas importantes, facilitando así información clave para el diagnóstico (14).

Además, fuera del ámbito práctico, la IA es capaz de realizar predicciones más precisas de supervivencia para ciertos procesos oncológicos que las clasificaciones actuales. Estos avances permiten que las interpretaciones sean más eficientes, exactas y apropiadas (13).

Justificación y objetivos

Justificación

Este Trabajo de Fin de Grado trata de explorar las posibilidades que ofrece la Inteligencia Artificial, como herramienta de nueva aparición, para facilitar el diagnóstico precoz de las enfermedades raras.

Las enfermedades raras suponen un potencial reto para los médicos debido a su baja prevalencia, lo cual limita el conocimiento y la experiencia clínica disponible. Esta carencia de familiaridad con estas patologías suele conllevar diagnósticos erróneos o

tardíos, teniendo un impacto negativo sobre la calidad de vida de los pacientes y las personas que los rodean.

Por consiguiente, es fundamental investigar qué puede aportar la IA a la resolución de estos desafíos. La capacidad de la IA de analizar enormes conjuntos de datos ofrece una nueva coyuntura en términos de precisión y celeridad para el diagnóstico y pronóstico de enfermedades raras. Además, la IA es capaz de identificar patrones complejos y establecer relaciones entre datos que podrían pasar inadvertidas para el ojo humano. Facilitar el diagnóstico tendría un impacto positivo sobre la calidad de vida de los pacientes, lo que a su vez podría mejorar los resultados y ofrecer atención médica más personalizada.

Por tanto, se hace necesario sintetizar el conocimiento disponible en materia de IAs aplicadas al diagnóstico de enfermedades raras, dado su rápido avance durante los últimos años y la previsible evolución al alza que experimentará en las próximas décadas.

Objetivo General

Analizar la posibilidad de aplicación de la Inteligencia Artificial en el diagnóstico precoz de enfermedades raras.

Objetivos específicos

- **OE 1:** Analizar si la disponibilidad de datos sobre enfermedades raras puede afectar al entrenamiento de Inteligencias Artificiales.
- **OE 2:** Analizar qué aspectos del diagnóstico precoz de enfermedades raras podrían beneficiarse del uso de IAs.
- **OE 3:** Analizar y discutir las consideraciones éticas que plantea la llegada de la IA al campo de la Medicina.

Materiales y métodos

Estrategia de búsqueda y criterios de selección

Para la realización de la presente revisión se han consultado diversas bases de datos relacionadas con el campo de la Medicina, como son PubMed, Cochrane y Medline (OvidSP).

Se inició la búsqueda en PubMed con la palabra clave “*artificial intelligence*”, obteniendo un total de 294.591 artículos. A continuación, se aplicaron los siguientes criterios de selección: que los documentos obtenidos fuesen revisiones, que incluyesen la palabra clave “*rare diseases*” y que a su vez estuvieran dentro de MeSH (*Medical Subject Headings*) y publicados en los últimos 5 años (2021-2025). Como resultado se consiguieron 31 revisiones, de las cuales seis fueron descartadas por no aportar una visión global del tema.

Se repitió la misma búsqueda en Cochrane, pero no se obtuvieron resultados relevantes para este trabajo. Por último, en Ovid como encabezamientos de materia se emplearon los términos “*artificial intelligence*” y “*rare diseases*”. Se seleccionó la casilla “centrar” para que estos términos fueran el punto principal de los resultados y se limitó la búsqueda a revisiones publicadas en los últimos 5 años. Esta búsqueda arrojó 21 artículos coincidentes con los anteriormente encontrados en PubMed.

Para completar el análisis, se hizo una búsqueda adicional en las mismas plataformas con el objetivo de proporcionar ejemplos concretos no identificados previamente.

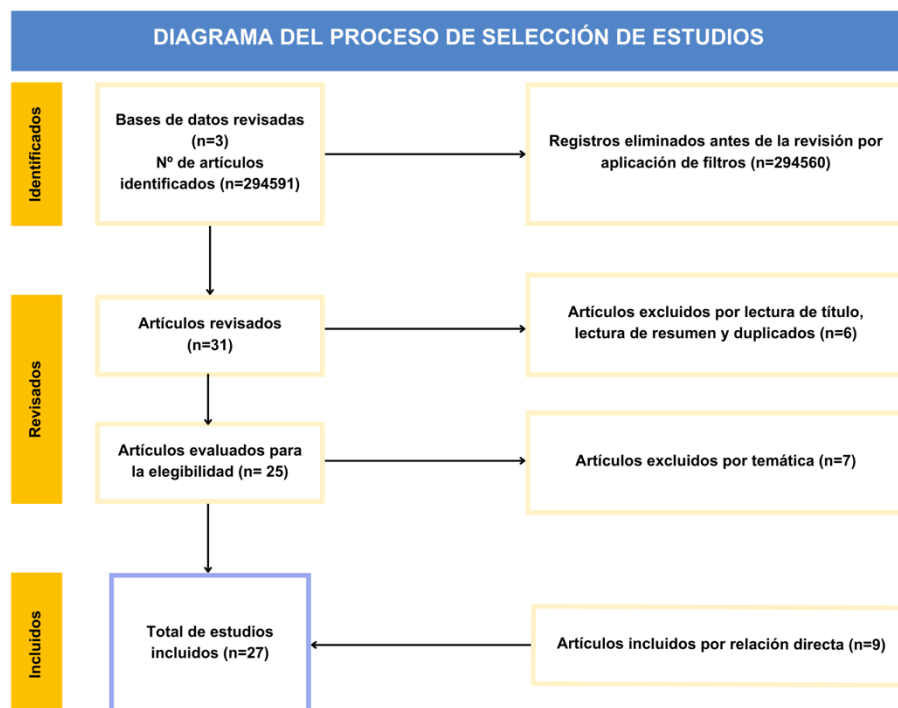


Figura 3: Diagrama del proceso de selección de estudios incluidos en el TFG. Elaboración propia.

Tras la lectura completa de cada una de las publicaciones anteriormente mencionadas, se seleccionaron un total de 18 artículos, para incluir en esta revisión. El motivo principal de descarte fue la falta de información relevante para responder a la pregunta sobre aplicaciones de la IA en el diagnóstico de enfermedades raras.

Así mismo, se han incluido 9 artículos que no cumplían con los criterios anteriores, cuya información proporcionada estaba directamente relacionada con los temas a tratar (Figura 3).

Por último, se consultó la Legislación Europea en materia de Inteligencia Artificial.

Evaluación de la calidad

Para obtener datos de calidad se han utilizado mayoritariamente artículos académicos situados en el primer cuartil (Q1). Además, se ha intentado buscar información lo más actualizada posible, por lo que la mayoría de los documentos que tratan información nueva en la revisión han sido publicados durante los últimos cinco años (Figura 4).

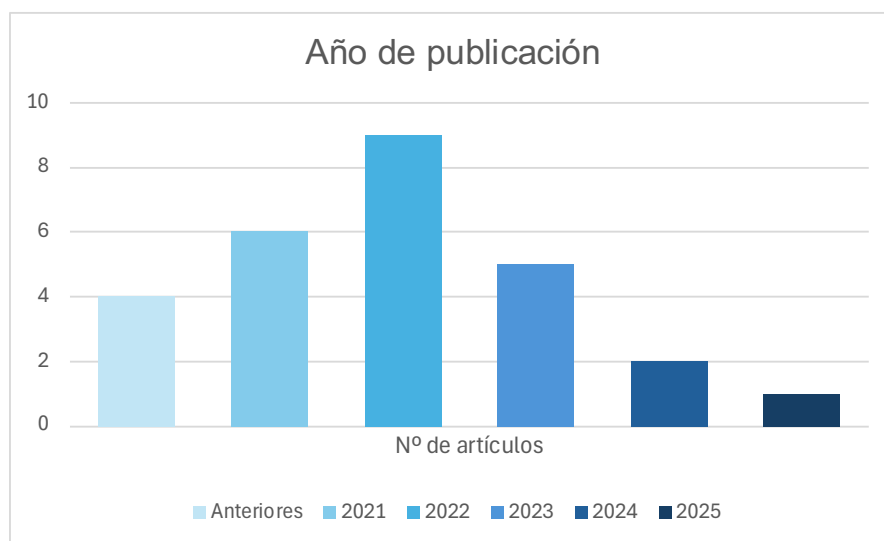


Figura 4: Gráfica de años de publicación. Elaboración propia.

Resultados

Cómo afecta la limitación de la disponibilidad de datos al entrenamiento de la IA

La primera limitación radica en el pequeño tamaño muestral. Existen dificultades para encontrar pacientes con enfermedades raras y los tamaños de muestra de los estudios suelen ser pequeños. Esta es la limitación afecta al 11% de las publicaciones sobre el mapeo de bases de datos para enfermedades raras. Esto limita la capacidad de extraer conclusiones generalizables con los datos recopilados (7).

Además, los algoritmos de machine learning corren el riesgo de heredar sesgos presentes en los datos de entrenamiento, lo que puede conllevar disparidades diagnósticas. De aquí surge la necesidad de desarrollar métodos para identificar y paliar estos sesgos (15). Este hecho, sumado a la reducida disponibilidad de datos de calidad, pone trabas al entrenamiento efectivo de los algoritmos de IA, problema que, sumado a que la calidad de los datos disponibles no es siempre la deseable, afecta a la precisión de los modelos (16).

En segundo lugar, cabe destacar el acceso limitado a los datos. Hay muchas fuentes de datos que son privadas o de pago, lo que dificulta el acceso con fines de investigación (17). En la investigación de las enfermedades raras es esencial que la información se comparta de manera internacional, para poder encontrar casos similares en el mayor número de pacientes posibles (7).

Normalmente, se utilizan datos estructurados que provienen de registros de salud electrónicos, lo que permite seleccionar variables fiables de alto valor predictivo positivo. Esto deja de lado los datos no estructurados (por ejemplo, imágenes), que a menudo contienen información clínica relevante, como la severidad de la enfermedad, su comportamiento o limitaciones que podrían ser factores pronósticos importantes que considerar en los análisis (18).

Como los datos de la historia natural proceden de la práctica clínica, este conjunto de datos puede hacer referencia a poblaciones menos uniformes que la que se selecciona para un ensayo clínico aleatorizado, lo que puede afectar a la recopilación e interpretación de los datos (17). Debido a las dificultades para encontrar muestras de tamaño suficiente en el caso de las enfermedades raras, puede ser interesante recolectar datos con métodos más avanzados, o combinar datos disponibles para crear nuevas variables que analizar (18).

[Análisis de aspectos del diagnóstico precoz de enfermedades raras que podrían beneficiarse del uso de IAs](#)

La IA es capaz de revisar e integrar conjuntos enormes de datos clínicos, genéticos y fenotípicos para elaborar listas de diagnósticos diferenciales relevantes, como es el caso de la *Human Phenotype Ontology* (HPO), que consiste en un vocabulario estandarizado de anomalías fenotípicas de enfermedades humanas (19).

Los modelos supervisados, como las redes neuronales, han demostrado su efectividad para la clasificación de enfermedades a partir de datos etiquetados. Por contraparte, los métodos no supervisados, como el *clustering* y los *autoencoders*, han sido útiles para detectar patrones ocultos en los datos no etiquetados (20), lo que es especialmente relevante dado que muchas enfermedades raras carecen de bases de datos extensas y bien definidas (17).

Por otro lado, a través de redes bayesianas se pueden identificar subtipos raros de enfermedades y relaciones de comorbilidad. En el contexto del diagnóstico, una red bayesiana integra información clínica, antecedentes y hallazgos del paciente para calcular la probabilidad de que presente una enfermedad o subtipo específico, incluso cuando los síntomas son poco frecuentes o se solapan con otras patologías. Esto se puede ver en el caso del asma, cuando tiene una presentación atípica, donde se sugiere que los ejes de su red de comorbilidad son la enfermedad pulmonar obstructiva crónica (EPOC), la insuficiencia respiratoria, la hipertensión, la aterosclerosis, y la gastritis y duodenitis, como indican Yong Z et al. en su análisis de costes y comorbilidad de la enfermedad (21).

Una de las principales ventajas de las enfermedades raras, desde el punto de vista del análisis de datos, radica en que alrededor del 80% poseen una base genética, mientras que las IAs tienen una demostrada capacidad para analizar secuencias de nucleótidos y aminoácidos. De esa forma, las IAs pueden identificar variantes patogénicas y correlacionarlas con fenotipos específicos como ya hacen, por ejemplo, PhenIX y Xrare (12).

Los diferentes algoritmos de IA pueden, en casos seleccionados, llegar a superar la capacidad humana para la identificación de patrones complejos, lo que se traduce en diagnósticos más precisos y fiables (22).

Por último, la capacidad de la IA de aprender a partir de grandes volúmenes de datos permite una identificación más veloz de hipotéticas enfermedades, minimizando el intervalo de tiempo desde la sospecha hasta el diagnóstico final (23).

Consideraciones éticas que plantea la llegada de la IA para el ejercicio de la profesión médica

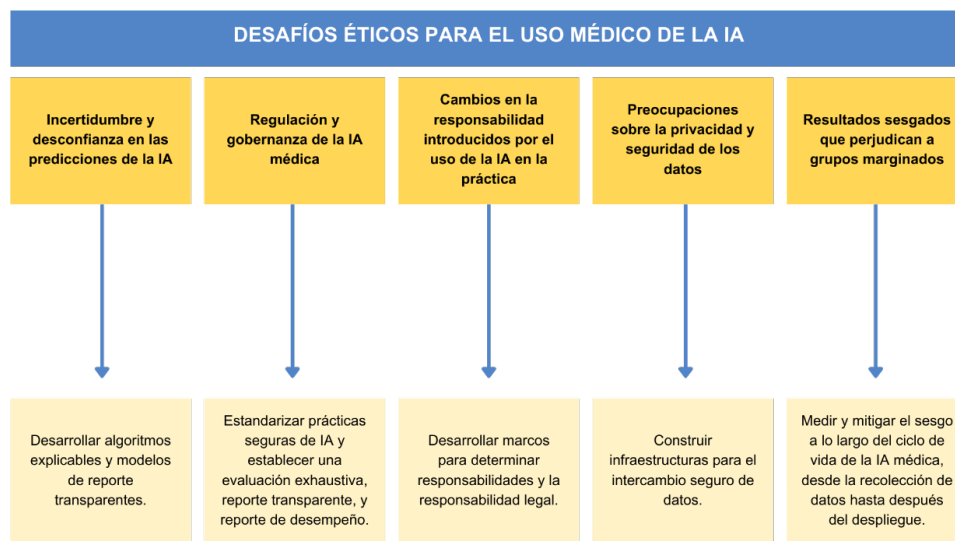


Figura 5: Desafíos éticos que plantea la llegada de la IA a la Medicina. Reproducido de (13).

Al tratarse de enfermedades raras, los riesgos de la utilización de la inteligencia artificial son relativamente similares a sus beneficios. A la hora del diagnóstico y la planificación terapéutica, es necesario recordar que la seguridad jurídica y la protección de datos, así como la funcionalidad (entendida como rapidez y veracidad diagnóstica) son los puntos clave del uso de la IA en medicina (13).

Dado que la responsabilidad sigue recayendo en el médico, es importante ofrecerle un uso seguro de las herramientas digitales dentro del marco legal. Esto implica la inclusión de nuevos programas en los sistemas sanitarios, que aseguren la encriptación y trazabilidad de los datos que se les proporcionen, y la garantía de que el sistema utilizado haya sido capaz de aprender tanto de resultados favorables como de los que no lo son, evitando así el sesgo de entrenamiento (24).

La Unión Europea es pionera en materia de legislación en lo que a IA se refiere. En el año 2021, la Comisión Europea presentó una propuesta de Ley de IA. En diciembre de 2023 se alcanzó un acuerdo definitivo sobre este texto, que ha sido aprobado en marzo del 2024, con intención de implantarse paulatinamente hasta el año 2026.

En cuanto a riesgos para la salud y seguridad, se plantea la existencia de sistemas de IA de alto riesgo, definidos porque pueden suponer perjuicio para la salud y la seguridad

o por la capacidad de ir contra los derechos fundamentales. Los riesgos son evaluados en base a una serie de criterios entre los que se encuentran la finalidad del sistema, el historial de incidentes previos y la vulnerabilidad de los afectados. Para mantener la seguridad y transparencia de estos sistemas de IA de alto riesgo, se requerirá que se mantengan registros e información técnica detallada sobre su desarrollo y funcionamiento en el transcurso de su vida útil. Incluirá algoritmos, datos, métodos de entrenamiento, prueba y validación (25).

Otra de las claves de regulación es la valoración de la creación de un espacio europeo de datos sanitarios que facilite el acceso de manera no discriminatoria. Esto permitirá el entrenamiento de algoritmos de manera transparente y que asegure la privacidad de los datos bajo gobernanza institucional. También será necesaria la vigilancia humana de los sistemas durante su ciclo vital (25).

El objetivo es minimizar los riesgos para la salud y evitar consecuencias negativas del uso o interpretación de la información generada.

Discusión

La posibilidad de utilizar la inteligencia artificial (IA) para diagnosticar enfermedades raras supone un avance significativo para la Medicina, ofreciendo potenciales soluciones para problemas históricos como el retraso diagnóstico, la falta de tratamiento y los errores clínicos (26).

Los resultados derivados de este trabajo señalan que los algoritmos de IA suponen una oportunidad para aumentar la precisión y velocidad diagnóstica de enfermedades raras (26).

Uno de los principales desafíos identificados en la revisión es el pequeño tamaño muestral asociado a las enfermedades raras. Su baja prevalencia y la dificultad del acceso a datos complican la recopilación, lo que restringe la capacidad para obtener conclusiones generalizables (17). Esto coincide con estudios previos, que subrayan la dificultad de trabajar con enfermedades raras debido a la limitada disponibilidad de datos clínicos (16).

Además, los algoritmos de IA corren el riesgo de heredar sesgos de los datos de entrenamiento, lo que puede resultar en diagnósticos erróneos (15). La calidad de los

datos también es un problema significativo: datos de baja calidad pueden reducir la precisión de los modelos, como se ha documentado en otras investigaciones (16).

El diagnóstico de enfermedades raras presenta un retraso promedio de hasta 8 años (7). La IA ofrece una solución potencial mediante la integración y análisis de grandes conjuntos de datos clínicos, genéticos y fenotípicos. Herramientas como el HPO demuestran cómo la IA es capaz de generar listas de diagnósticos diferenciales (19).

Los modelos supervisados, como las redes neuronales, han mostrado eficacia en la clasificación de enfermedades a partir de datos etiquetados, mientras que los métodos no supervisados, como el *clustering* y los *autoencoders*, son útiles para detectar patrones ocultos en datos no etiquetados, lo que resulta útil en el caso de las enfermedades raras (20).

Dado que la base de aproximadamente el 80% de las enfermedades es genética, la capacidad de la IA para analizar secuencias de nucleótidos y aminoácidos presenta ventajas. Algoritmos como PhenIX y Xrare ya están identificando variantes patogénicas y correlacionándolas con fenotipos específicos, demostrando el potencial de la IA para superar la capacidad humana en la identificación de patrones complejos (12) (22).

La seguridad jurídica y la protección de datos son esenciales en el uso de IA en medicina. La responsabilidad médica implica que los sistemas de IA deben ser seguros y funcionar dentro de un marco legal claro. Es necesario asegurar la trazabilidad y seguridad de los datos utilizados en estos sistemas.

La Unión Europea ha tomado la delantera en la regulación de la IA. La Ley de IA establece criterios para la seguridad y transparencia de los sistemas de IA de alto riesgo, incluyendo la necesidad de mantener registros detallados sobre su desarrollo y funcionamiento. Esto asegura que los algoritmos sean entrenados de manera transparente y que se respete la privacidad de los datos (25).

Direcciones futuras

Para superar las barreras actuales y maximizar el potencial de la IA en el diagnóstico de enfermedades raras, se realizan varias propuestas para enfocar el futuro.

Es esencial crear iniciativas para compilar y compartir datos de manera segura y privada entre instituciones, y así poder proporcionar conjuntos de datos más robustos para el entrenamiento de algoritmos (17).

Otra propuesta es el desarrollo de modelos híbridos, debido a que la combinación de aprendizaje supervisado y no supervisado, así como la integración de diferentes tipos de datos (genómicos, clínicos, de imagen), puede mejorar la fiabilidad y aplicabilidad de los diagnósticos (27).

También es necesario aumentar la interpretabilidad mediante el desarrollo de técnicas y la formación de profesionales, con el objetivo de mejorar la comprensión acerca de cómo los modelos de IA pueden incrementar la confianza de los médicos y pacientes en estas tecnologías (18).

Para finalizar, se necesita establecer políticas transparentes y concisas, que guíen el uso de la IA en el ámbito sanitario, asegurando que se vele por la integridad de los principios éticos y legales (24).

Conclusión

Como conclusión, la revisión llevada a cabo ha permitido comprobar que la falta de datos podría ser un factor limitante a la hora de entrenar IAs para el diagnóstico de enfermedades raras. La capacidad de las IAs de procesar fácilmente datos de secuenciación masiva facilitaría dicha tarea, ya que la mayoría de las enfermedades raras tienen una base genética. Por último, aunque existen esfuerzos en materia de ética y seguridad, la legislación aún está en desarrollo en la mayoría de los países, siendo la UE pionera en la implantación de leyes internacionales que regulen la utilización de la IA.

Bibliografía

1. Rare diseases - European Commission [Internet]. [citado 4 de mayo de 2025]. Disponible en: https://research-and-innovation.ec.europa.eu/knowledge-publications-tools-and-data/publications/all-publications/rare-diseases_en
2. International Classification of Diseases (ICD) [Internet]. [citado 4 de mayo de 2025]. Disponible en: <https://www.who.int/standards/classifications/classification-of-diseases>
3. Groft SC, Posada M, Taruscio D. Progress, challenges and global approaches to rare diseases. *Acta Paediatr.* 2021;110(10):2711-6.
4. Garbade SF, Zielonka M, Komatsuzaki S, Kölker S, Hoffmann GF, Hinderhofer K, et al. Quantitative retrospective natural history modeling for orphan drug development. *J Inherit Metab Dis.* 1 de enero de 2021;44(1):99-109.
5. Zampatti S, Peconi C, Megalizzi D, Calvino G, Trastulli G, Cascella R, et al. Innovations in Medicine: Exploring ChatGPT's Impact on Rare Disorder Management. *Genes.* abril de 2024;15(4):421.
6. Hurvitz N, Azmanov H, Kesler A, Ilan Y. Establishing a second-generation artificial intelligence-based system for improving diagnosis, treatment, and monitoring of patients with rare diseases. *Eur J Hum Genet.* octubre de 2021;29(10):1485-90.
7. Raycheva R, Kostadinov K, Mitova E, Bogoeva N, Iskrov G, Stefanov G, et al. Challenges in mapping European rare disease databases, relevant for ML-based screening technologies in terms of organizational, FAIR and legal principles: scoping review. *Front Public Health.* 2023;11:1214766.
8. Rare diseases, common challenges. *Nat Genet.* marzo de 2022;54(3):215-215.
9. Germain DP, Gruson D, Malcles M, Garcelon N. Applying artificial intelligence to rare diseases: a literature review highlighting lessons from Fabry disease. *Orphanet J Rare Dis.* 17 de abril de 2025;20:186.
10. Sharma N, Sharma R, Jindal N. Machine Learning and Deep Learning Applications-A Vision. *Glob Transit Proc.* junio de 2021;2(1):24-8.
11. Garcelon N, Burgun A, Salomon R, Neuraz A. Electronic health records for the diagnosis of rare diseases. *Kidney Int.* abril de 2020;97(4):676-86.
12. Krawitz PM. Künstliche Intelligenz bei der Diagnose Seltener Erkrankungen: die Entwicklung der Phänotyp-Analyse. *Bundesgesundheitsblatt - Gesundheitsforschung - Gesundheitsschutz.* 1 de noviembre de 2022;65(11):1159-63.
13. Rajpurkar P, Chen E, Banerjee O, Topol EJ. AI in health and medicine. *Nat Med.* enero de 2022;28(1):31-8.
14. Maier K, Zaniolo L, Marques O. Image quality issues in teledermatology: A comparative analysis of artificial intelligence solutions. *J Am Acad Dermatol.* julio de 2022;87(1):240-2.
15. Hasani N, Farhadi F, Morris MA, Nikpanah M, Rahmim A, Xu Y, et al. Artificial Intelligence in Medical Imaging and its Impact on the Rare Disease Community: Threats, Challenges and Opportunities. *PET Clin.* 1 de enero de 2022;17(1):13-29.

16. Deep learning for rare disease: A scoping review - ScienceDirect [Internet]. [citado 2 de mayo de 2025]. Disponible en: <https://www.sciencedirect.com/science/article/pii/S1532046422002325?via%3Dihub>
17. Liu Z, Zhu L, Roberts R, Tong W. Toward Clinical Implementation of Next-Generation Sequencing-Based Genetic Testing in Rare Diseases: Where Are We? *Trends Genet TIG*. noviembre de 2019;35(11):852-67.
18. Liu J, Barrett JS, Leonardi ET, Lee L, Roychoudhury S, Chen Y, et al. Natural History and Real-World Data in Rare Diseases: Applications, Limitations, and Future Perspectives. *J Clin Pharmacol*. diciembre de 2022;62(Suppl 2):S38-55.
19. Haimel M, Pazmandi J, Heredia RJ, Dmytrus J, Bal SK, Zoghi S, et al. Curation and expansion of Human Phenotype Ontology for defined groups of inborn errors of immunity. *J Allergy Clin Immunol*. enero de 2022;149(1):369-78.
20. Pratella D, Ait-El-Mkadem Saadi S, Bannwarth S, Paquis-Fluckinger V, Bottini S. A Survey of Autoencoder Algorithms to Pave the Diagnosis of Rare Diseases. *Int J Mol Sci*. enero de 2021;22(19):10891.
21. Yong Z, Luo L, Gu Y, Li C. Bayesian Comorbidity Network and Cost Analysis for Asthma. *IEEE J Biomed Health Inform*. septiembre de 2022;26(9):4714-24.
22. Schaefer J, Lehne M, Schepers J, Prasser F, Thun S. The use of machine learning in rare diseases: a scoping review. *Orphanet J Rare Dis*. 9 de junio de 2020;15(1):145.
23. Wojtara M, Rana E, Rahman T, Khanna P, Singh H. Artificial intelligence in rare disease diagnosis and treatment. *Clin Transl Sci*. 30 de agosto de 2023;16(11):2106-11.
24. Völkel L, Wagner AD. Mit künstlicher Intelligenz schneller zur Diagnose seltener Erkrankungen – ein Gebot der Ethik, Ökonomie und Lebensqualität. *Inn Med Heidelb Ger*. 2023;64(11):1033-40.
25. Proposal for a Regulation laying down harmonised rules on artificial intelligence | Shaping Europe's digital future [Internet]. [citado 4 de mayo de 2025]. Disponible en: <https://digital-strategy.ec.europa.eu/en/library/proposal-regulation-laying-down-harmonised-rules-artificial-intelligence>
26. Yang TH, Liao ZY, Yu YH, Hsia M. RDDDL: A systematic ensemble pipeline tool that streamlines balancing training schemes to reduce the effects of data imbalance in rare-disease-related deep-learning applications. *Comput Biol Chem*. 1 de octubre de 2023;106:107929.
27. Skweres-Kuchta M, Czerska I, Szaruga E. Literature Review on Health Emigration in Rare Diseases-A Machine Learning Perspective. *Int J Environ Res Public Health*. 30 de enero de 2023;20(3):2483.

Anexos

DEBILIDADES Y FORTALEZAS DEL USO DE INTELIGENCIA ARTIFICIAL PARA EL DIAGNÓSTICO DE ENFERMEDADES RARAS



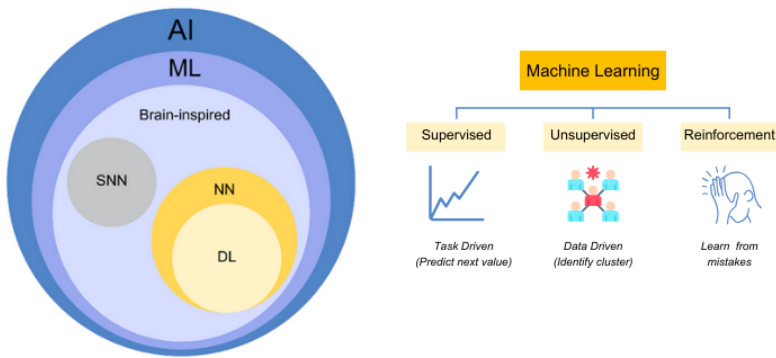
Universidad de Valladolid

Autor: Ana Bestard Cassinello
Tutores: Diego Sánchez Romero, Sergio Díez Hermano

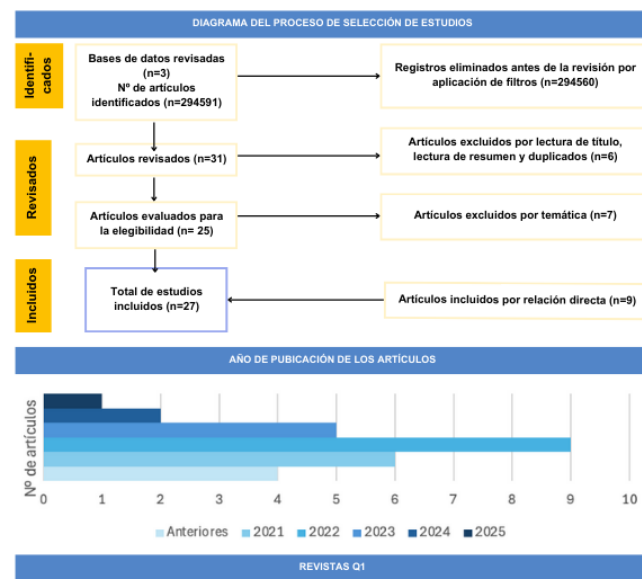
Resumen: Las enfermedades raras suponen más de 400 millones de casos en el mundo. Desde los años 90, han ganado peso para la salud pública. La IA, puede mejorar los diagnósticos mediante el análisis de datos guiado y por el aprendizaje automático. El objetivo general es analizar la posible aplicación de la IA en este campo y la limitación que supone la reducida disponibilidad de datos, los beneficios potenciales del diagnóstico precoz y las consideraciones éticas que supone la llegada de la IA a la ciencia médica. Tras la búsqueda, la selección final consistió en 27 artículos. Se priorizó que fueran recientes, de 2020 en adelante, y de primer cuartil. También se revisó la Legislación Europea sobre IA. Esta revisión ha permitido comprobar que IA puede mejorar el diagnóstico de enfermedades raras, reduciendo tiempos y aumentando la precisión. Aunque enfrenta desafíos relacionados con la disponibilidad y calidad de datos, sesgos, interpretabilidad y necesidad de regulación.

INTRODUCCIÓN

- **Prevalencia:** 3-6% de la población (400 millones de personas).
- **7000** enfermedades raras conocidas, **80% con origen genético**.
- 8 años de media de **retraso diagnóstico**.
- IA: **análisis masivo de datos e identificación de patrones** supone posible mejora en diagnóstico.



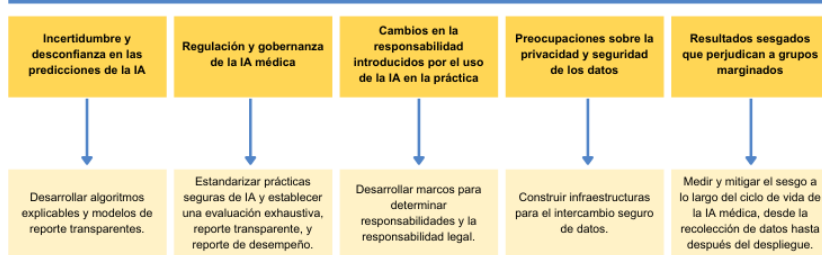
METODOLOGÍA



RESULTADOS

- La **limitada disponibilidad de datos** sobre enfermedades raras, dificulta el entrenamiento de IA.
- La IA puede **mejorar el diagnóstico precoz** integrando datos clínicos, genéticos y fenotípicos, y analizando imágenes con algoritmos de aprendizaje profundo.
- Puede identificar **subtipos raros y comorbilidades** mediante redes bayesianas.
- Sin embargo, el uso de IA plantea **desafíos éticos** relacionados con la seguridad jurídica y la protección de datos.
- La **Unión Europea** aborda esto con legislación específica que incluye la creación de un espacio europeo de datos sanitarios, asegurando el acceso no discriminatorio, la supervisión humana y la transparencia de los algoritmos.

DESAFÍOS ÉTICOS PARA EL USO MÉDICO DE LA IA



DISCUSIÓN

- La IA puede **mejorar el diagnóstico** de enfermedades raras, reduciendo retrasos y errores clínicos.
- Desafíos: pequeño **tamaño muestral** y riesgo de **sesgos** en los datos de entrenamiento.
- La IA identifica **variantes genéticas y patrones complejos**.
- Esencial: seguridad jurídica y **protección de datos**.
- **Futuras investigaciones:** compilación segura de datos, desarrollo de modelos híbridos y mejora de la interpretabilidad de los modelos de IA.

CONCLUSIÓN

- La **falta de datos** podría **limitar el entrenamiento de IAs** para el diagnóstico de enfermedades raras.
- La capacidad de las IAs de procesar datos de **secuenciación masiva** facilita esta tarea.
- Existen esfuerzos en materia de **ética y seguridad**, pero la **legislación está en desarrollo** en la mayoría de países.
- La **UE es pionera** en la implantación de leyes que regulen el uso de la IA.



Accede al texto completo



Accede a la bibliografía