



# INTELIGENCIA ARTIFICIAL APLICADA AL DIAGNÓSTICO DE ENFERMEDADES TROPICALES DESATENDIDAS

La implementación de la inteligencia artificial en la investigación biomédica está permitiendo avanzar en el desarrollo de un diagnóstico cada vez más rápido, completo y automatizado de enfermedades complejas como el cáncer. Pero su aplicación también puede ser muy beneficiosa en la lucha contra otras patologías cuya investigación no recibe tanta atención ni financiación, como las denominadas por la OMS como Enfermedades Tropicales Desatendidas.

**ANA TORRES,**

Centro Colaborador de la OMS para Leishmaniasis, Centro Nacional de Microbiología, CIBER de Enfermedades Infecciosas, Instituto de Salud Carlos III.

**LORENA BERNARDO,**

Centro Colaborador de la OMS para Leishmaniasis, Centro Nacional de Microbiología, CIBER de Enfermedades Infecciosas, Instituto de Salud Carlos III.

**MARINA LOZANO,**

Centro Nacional de Microbiología, Instituto de Salud Carlos III.

**JOSE CARLOS SOLANA,**

Centro Colaborador de la OMS para Leishmaniasis, Centro Nacional de Microbiología, CIBER de Enfermedades Infecciosas, Instituto de Salud Carlos III.

## La leishmaniasis dérmica post-kala-azar, una enfermedad olvidada

La leishmaniasis afecta principalmente a las personas pobres en países tropicales y subtropicales de todo el mundo, generando a su vez más pobreza. En la Cuenca Mediterránea es una enfermedad endémica, siendo **España el país europeo con más casos reportados**. Se transmite por insectos flebótomos (o moscas de la arena)

que se alimentan de sangre e introducen durante la picadura parásitos *Leishmania* que infectan las células del paciente. Las dos principales manifestaciones clínicas de la enfermedad son la cutánea, que produce lesiones en la piel, y la visceral, en la que los parásitos invaden órganos internos como el hígado y el bazo y producen la muerte del paciente en ausencia de un tratamiento. Una forma poco común de esta enfermedad es la **leishmaniasis dérmica post-kala-azar** (o PKDL, por sus siglas

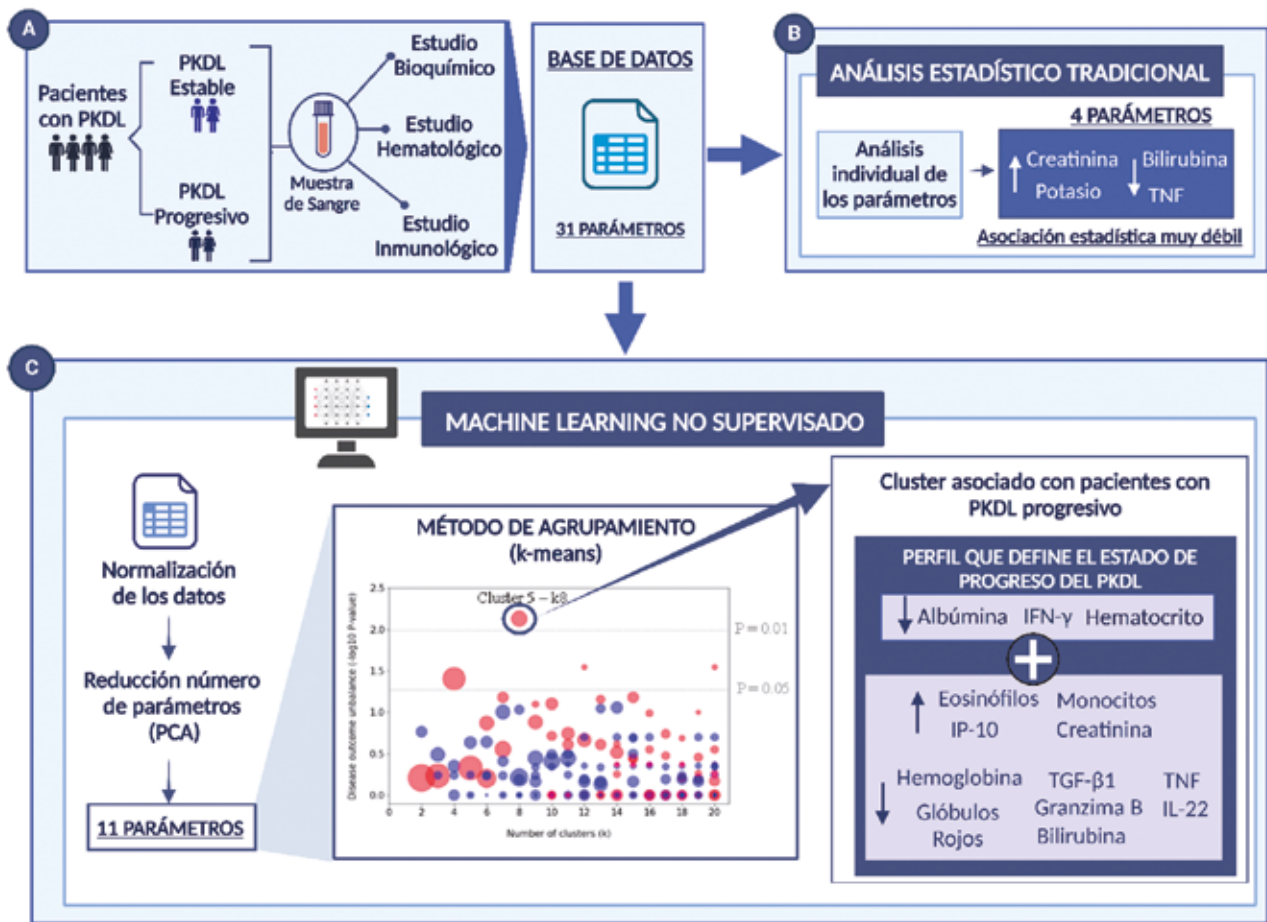


Figura 1. Flujo de trabajo para el uso de IA en un ensayo clínico de leishmaniasis.

en inglés) que se produce en algunas regiones del subcontinente Indio y en África Oriental tiempo después de que los pacientes se hayan curado de una leishmaniasis visceral<sup>1</sup>. Aparecen lesiones en la piel en forma de nódulos que no solo son muy estigmatizantes para los pacientes, sino que estos además pueden servir de **reservorios de la enfermedad**, pudiendo transmitir los parásitos y generar nuevos focos de infección. El menor número de casos de PKDL comparado con otras manifestaciones de la leishmaniasis, sumado a que el PKDL no supone un riesgo a la vida del paciente hace que **el estudio y control del PKDL esté aún más desatendido**. Es necesario, por tanto, **identificar biomarcadores** que informen sobre el estado de estos pacientes e identifiquen una mala evolución de su enfermedad, y así iniciar cuanto antes un tratamiento adecuado. Sin embargo, apenas existen estudios clínicos al respecto, y los datos obtenidos de ellos no han permitido realizar grandes avances. Además, y de manera general, resulta cada vez es más difícil poner en marcha ensayos clínicos en los que participe el número de pacientes

necesario como para obtener conclusiones relevantes. En este contexto, **las herramientas basadas en inteligencia artificial (IA)** pueden ser la clave que nos permita optimizar los ensayos clínicos, sacar más provecho de los datos que ya existen y ayudar a identificar biomarcadores de interés.

### ¿Puede la IA encontrar biomarcadores de PKDL en un análisis de sangre?

Como prueba de concepto para evaluar la utilidad de la IA en el diagnóstico de PKDL, se han analizado los datos obtenidos de un **análisis de sangre** de pacientes con PKDL incluidos en un ensayo clínico en el que ha participado el Centro Colaborador de la OMS para Leishmaniasis del Instituto de Salud Carlos III. Estos pacientes con PKDL pueden dividirse en dos grupos, uno con una situación controlada y sin cambios en los últimos meses (**PKDL estable**) y otro en el que la enfermedad había empeorado (**PKDL progresivo**). En este contexto, el objetivo fue

encontrar entre los datos obtenidos algún parámetro que pudiera ser utilizado como un **biomarcador que pronostique un agravamiento de la enfermedad**. Entre los 31 parámetros analizados hay resultados bioquímicos comunes, como la concentración de creatinina en plasma y los niveles de las enzimas hepáticas, la concentración de hemoglobina o el número de leucocitos (Figura 1A). También se incluyen resultados de los análisis inmunológicos en los que se mide la producción de citoquinas por las células sanguíneas cuando se estimulan con los antígenos de *Leishmania* (respuesta inmunitaria celular).

En un **análisis estadístico tradicional**, cada uno de estos parámetros se analizaría de forma **individual** para ver si existe una asociación significativa con aquellos individuos que han empeorado. Sin embargo, estos análisis solo han permitido observar una **asociación estadísticamente significativa muy débil** (cerca de  $p = 0.05$ ) con el aumento de los niveles de creatinina, potasio, o con una bajada de los niveles de bilirrubina y del factor de necrosis tumoral (TNF) (Figura 1B).

La limitada capacidad del análisis convencional de los datos puede superarse con herramientas basadas en IA capaces de sacar a la luz información valiosa que subyace en las relaciones existentes entre todos los parámetros. En este caso se utilizaron todos los datos recogidos en el análisis de sangre para llevar a cabo un aprendizaje automatizado no supervisado (o *machine learning no supervisado*)<sup>2</sup> (Figura 1C). Esta aproximación permite detectar agrupaciones naturales de pacientes sin dar ningún tipo de información al algoritmo sobre su estado de PKDL (estable o progresivo). Para realizar este tipo de análisis multi-paramétricos es necesario, en primer lugar, **preparar los datos** para integrarlos en la misma escala y así evitar un sesgo por una mayor influencia de algunos parámetros o por individuos que se salgan de la norma general (normalización). A continuación, es necesario reducir el número de parámetros mediante un **análisis de componentes principales** (PCA, por sus siglas en inglés) que ayude a condensar la información de múltiples variables en un número reducido de componentes principales que expliquen la mayor variabilidad de los datos. Este es un paso crucial para capturar las correlaciones más importantes y facilitar la aplicación de los algoritmos de agrupamiento. En este caso se consiguió seleccionar **11 parámetros** (de los 31 iniciales) que acumulan más del 90% de la varianza de los datos y cuyas correlaciones deben explicar las diferencias que existen entre todos los pacientes.

Después de este pretratamiento, los datos ya se pueden aplicar a un método de **agrupamiento para capturar correlaciones complejas** entre el estado de la enfermedad (estable vs progresivo) y los 11 parámetros seleccionados. El método k-means es un algoritmo de agrupamiento que divide un conjunto de datos en un número específico

**Una aproximación basada en pruebas de laboratorio convencionales y técnicas avanzadas de inteligencia artificial es capaz de identificar variaciones específicas en factores bioquímicos, hematológicos e inmunológicos que se relacionan con el estado de una enfermedad, en este caso el progreso del PKDL.**

de grupos o clústeres que contienen elementos similares entre sí, pero diferentes de los demás, y que permite encontrar patrones de correlación que no son obvios o detectables con análisis convencionales. Uno de estos clústeres se asocia de forma muy significativa ( $p = 0.0074$ ) con los pacientes en los que su PKDL había empeorado (Figura 1C). La comparativa de los parámetros de los pacientes de este clúster con los del resto de grupos muestra que tienen un menor hematocrito, niveles más bajos de albúmina y una menor producción de IFN- $\gamma$ . También comparten un significativo aumento en el porcentaje de eosinófilos y monocitos, y en los niveles de creatinina e IP-10, a la vez que una disminución de los valores de hemoglobina, IL-22, TNF, glóbulos rojos, granzima B, TGF- $\beta$  y bilirrubina. Por lo tanto, esta es la combinación de parámetros que resulta en un perfil que define el estado de progreso del PKDL.

En definitiva, una aproximación basada en pruebas de laboratorio convencionales y técnicas avanzadas de inteligencia artificial es capaz de identificar variaciones específicas en factores bioquímicos, hematológicos e inmunológicos que se relacionan con el estado de una enfermedad, en este caso el progreso del PKDL. Serán necesarios más estudios para poder introducir esta metodología en la práctica clínica, pero este esquema de trabajo basado en IA puede ya aplicarse a poblaciones más grandes afectadas de PKDL, con el objetivo de entrenar un algoritmo de *machine learning supervisado* que pueda en el futuro predecir de forma precisa el estado de PKDL o realizar un diagnóstico temprano, además de ayudarnos a comprender mejor esta enfermedad.

### Referencias

1. Zijlstra EE, Musa AM, Khalil EA, el-Hassan IM, el-Hassan AM. Post-kala-azar dermal leishmaniasis. *Lancet Infect Dis*. 2003;3(2):87-98.
2. Eckhardt CM, Madjarova SJ, Williams RJ, Ollivier M, Karlsson J, Pareek A, et al. Unsupervised machine learning methods and emerging applications in healthcare. *Knee Surg Sports Traumatol Arthrosc*. 2023;31(2):376-81