

EVIDENCIAS EN PEDIATRÍA

Toma de decisiones clínicas basadas en las mejores pruebas científicas
www.evidenciasenpediatria.es

Artículos Valorados Críticamente

Aprendizaje automático para identificar niños con fiebre con riesgo de presentar enfermedad de Kawasaki

Molina Arias M¹, Ortega Páez E²

¹Servicio de Gastroenterología. Hospital Infantil Universitario La Paz. Madrid. España.

²UGC Góngora. Distrito Granada-Metropolitano. Granada. España.

Correspondencia: Manuel Molina Arias: mma1961@gmail.com

Palabras clave en español: aprendizaje automático; enfermedad de Kawasaki; técnicas y procedimientos diagnósticos.

Palabras clave en inglés: machine learning; mucocutaneous lymph node syndrome; diagnostic techniques and procedures.

Fecha de recepción: 8 de noviembre de 2023 • **Fecha de aceptación:** 5 de diciembre de 2023

Fecha de publicación del artículo: 10 de enero de 2024

Evid Pediatr. 2024;20:3.

CÓMO CITAR ESTE ARTÍCULO

Molina Arias M, Ortega Páez E. Aprendizaje automático para identificar niños con fiebre con riesgo de presentar enfermedad de Kawasaki. Evid Pediatr. 2024;20:3.

Para recibir Evidencias en Pediatría en su correo electrónico debe darse de alta en nuestro boletín de novedades en <http://www.evidenciasenpediatria.es>

Este artículo está disponible en: <http://www.evidenciasenpediatria.es/EnlaceArticulo?ref=2024;20:3>.

©2005-24 • ISSN: 1885-7388

Aprendizaje automático para identificar niños con fiebre con riesgo de presentar enfermedad de Kawasaki

Molina Arias M¹, Ortega Páez E²

¹Servicio de Gastroenterología. Hospital Infantil Universitario La Paz. Madrid. España.

²UGC Góngora. Distrito Granada-Metropolitano. Granada. España.

Correspondencia: Manuel Molina Arias: mma1961@gmail.com

Artículo original: Tsai CM, Lin CR, Kuo HC, Cheng FJ, Yu HR, Hung TC, et al. Use of machine learning to differentiate children with Kawasaki disease from other febrile children in a Pediatric Emergency Department. JAMA. 2023;6:e237489.

Resumen

Conclusiones de los autores del estudio: el estudio sugiere que los resultados de las pruebas objetivas de laboratorio tienen el potencial de predecir la enfermedad de Kawasaki. El aprendizaje automático con XGBoost puede ayudar a los clínicos a diferenciar a los pacientes con enfermedad de Kawasaki de otros pacientes febriles en los servicios de urgencias pediátricas, con una excelente sensibilidad, especificidad y precisión.

Comentario de los revisores: aunque el modelo presentado tiene buena potencia para identificar pacientes con riesgo de enfermedad de Kawasaki, es preciso realizar una validación externa del mismo en poblaciones más parecidas a la nuestra antes de poder recomendar su utilización.

Palabras clave: aprendizaje automático; enfermedad de Kawasaki; técnicas y procedimientos diagnósticos.

Machine learning to identify febrile children at risk of Kawasaki disease

Authors' conclusions: the study suggests that the results of objective laboratory tests have the potential to predict Kawasaki disease. Machine learning with XGBoost can help clinicians differentiate Kawasaki disease patients from other febrile patients in pediatric emergency departments with excellent sensitivity, specificity, and accuracy.

Reviewers' commentary: although the model presented has power to identify patients at risk of Kawasaki disease, it must be externally validated in populations more similar to ours before its use can be recommended.

Key words: machine learning; mucocutaneous lymph node syndrome; diagnostic techniques and procedures.

RESUMEN ESTRUCTURADO

Objetivo: desarrollar un modelo predictivo mediante aprendizaje automático con parámetros objetivos para diferenciar a los pacientes con enfermedad de Kawasaki (EK) de otros pacientes febriles.

Diseño: estudio de pruebas diagnósticas de casos y controles.

Emplazamiento: cuatro centros hospitalarios de Taiwan.

Población de estudio: durante un periodo de 10 años, se incluyeron en total 107 062 pacientes con fiebre a los que se les habían realizado pruebas diagnósticas que fueron el 22% de los que acudieron a urgencias por fiebre. Se seleccionaron 1158 casos de EK completa o incompleta conforme a los criterios de la American Heart Association (AHA) y 105 904 controles menores de 5 años con fiebre (no se especifica duración) que acudieron al mismo tiempo que los casos al centro hospitalario. Se excluyeron 18 casos y 32 405 controles por tener más de cinco años de edad. La muestra final fue de 1142 casos y 73 499 controles.

Evaluación de la prueba diagnóstica: las variables seleccionadas se recogieron de las historias clínicas electrónicas y se tuvieron en cuenta las notas clínicas para confirmar el diagnóstico de EK. Las variables incluidas en los dos grupos fueron: demográficas (edad, género), hemograma con recuento diferencial de células, proteína C reactiva (PCR), aspartato aminotransferasa (AST) y alanina aminotransferasa (ALT), sedimento urinario no recogido de forma estéril, codificado como variable dicotómica “Sí/No” según la existencia de piuria, medida como estearasa leucocitaria y como continua según el recuento de leucocitos. La velocidad de sedimentación globular (VSG) y el nivel de albúmina se excluyeron por falta de muestra. Los datos faltantes se trataron con relleno de la media del grupo cuando faltaban 1 o 2 resultados y se excluyeron si faltaban 3 o más resultados.

Medición del resultado: se realizaron pruebas t y exacta de Fisher para variables continuas o categóricas, respectivamente, y análisis de regresión logística binaria univariante y multivariante para identificar los factores de riesgo significativos para incluirlos en el modelo predictivo final. Se elaboró un algoritmo de aprendizaje automático supervisado. El algoritmo elegido fue Xtreme Gradient Boosting (XGBoost), basado en el ensamblaje de árboles de decisión y con una técnica de optimización de gradiente descendente. Este algoritmo proporciona buenos resultados cuando el etiquetado de las variables a incluir en el modelo está desequilibrado, como en el caso del conjunto de los pacientes con EK, y es capaz de capturar relaciones no lineales y realizar interacciones entre los datos. Se hizo previamente el conjunto de entrenamiento con una muestra aleatoria del 80% de los casos y el de validación con el 20% restante, sin diferencias significativas entre los dos.

Resultados principales: la media de edad fue de 1,1 años (desviación estándar [DE]: 0,8) y de 1,6 años (DE: 1,4) en los casos de EK y en los controles, respectivamente, sin diferencias estadísticas significativas. Hubo un predominio de varones en el grupo EK (60,2% frente a 56,4% en los controles, odds ratio [OR]: 1,79; intervalo de confianza al 95% [IC 95]: 1,55 a 2,06), que eran 0,6 años (IC 95: -0,5 a -0,6) más jóvenes. De las 5 variables incluidas en el modelo final, las que presentaron valores de importancia más altos, fueron: piuria (coeficiente de importancia 18,01), recuento de leucocitos en orina (12,63), nivel de ALT (7,65), nivel de PCR (7,14) y porcentaje de eosinófilos en sangre (4,97). El modelo de validación se hizo con el 20% de todos los datos del modelo (228 niños con EK, 14 700 controles). El modelo final, tras probar varios umbrales, obtuvo una sensibilidad (S): 92,5%, especificidad (E): 97,3%, valor predictivo positivo (VPP): 34,5%, valor predictivo negativo (VPN): 99,9%, cociente de probabilidad positivo (CPP): 34, cociente de probabilidad negativo (CPN): 0,08, área bajo la curva ROC (ABC): 0,98 (IC 95: 0,974 a 0,987). El modelo se probó en nuevos diagnósticos clasificando correctamente al 92,1% como EK.

Conclusión: este estudio diagnóstico sugiere que los resultados objetivos de las pruebas de laboratorio tienen potencial como parámetros predictivos de la EK. El modelo de predicción establecido mediante el enfoque de aprendizaje automático con XGBoost podría ayudar a los médicos de primera línea a detectar posibles pacientes con EK entre los niños febriles de un servicio de urgencias.

Conflicto de intereses: ninguno conocido.

Fuente de financiación: ninguna conocida.

COMENTARIO CRÍTICO

Justificación: el diagnóstico de la enfermedad de Kawasaki se basa en la presencia de signos clínicos que, además de precisar una valoración subjetiva, pueden presentarse de forma incompleta o asincrónica, lo que puede dificultar el diagnóstico precoz, sobre todo de las formas incompletas¹.

El abordaje diagnóstico con un método de aprendizaje automático puede permitir basar el diagnóstico en datos objetivos, posibilitándose el tratamiento precoz en una mayor proporción de estos pacientes.

Validez o rigor científico: existe una definición clara del escenario de aplicación del algoritmo diagnóstico, con definición de la población utilizada, las variables incluidas en el modelo (podría haber sido útil introducir como variable adicional el momento de realizar la analítica) y el patrón de referencia utilizado (criterios diagnósticos de la American Heart Association). La muestra de pacientes parece representativa de la enfermedad y se realiza de forma ciega e independiente la comparación entre la prueba diagnóstica y el patrón de referencia.

La prueba diagnóstica emplea un algoritmo de clasificación validado de aprendizaje automático supervisado (XGBoost) basado en árboles de decisión. El principal problema es el desequilibrio entre las dos categorías de la variable de resultado (1,5% de diagnóstico positivo), lo cual trata de corregirse mediante la aplicación de técnicas específicas de este algoritmo y la elección de una función de coste más adecuada para clases desbalanceadas.

No se realiza preprocesamiento de las variables (salvo la creación de variables indicadoras para las categóricas), ya que no es imprescindible debido a la naturaleza del algoritmo. El análisis de los datos se realiza de forma correcta, utilizando métricas adecuadas para clases desbalanceadas, como el área bajo la curva ROC.

Por último, aunque se realiza una división de los participantes que permite la validación interna de los resultados, no existe validación externa del modelo diagnóstico elaborado, lo que

cobra importancia debido al riesgo de sobreajuste de estos modelos.

Importancia clínica: con un ABC de 0,98 y un CPP de 34, el modelo muestra buena potencia para discriminar los casos de EK, sobre todo teniendo en cuenta la baja prevalencia de la enfermedad (0,01, 1158 de los 107 062 pacientes incluidos). Con los datos de S y E conseguidos, la probabilidad de padecer la enfermedad aumenta desde 0,01 hasta 0,25 cuando el modelo da un resultado positivo* (pensemos que un modelo que siempre predijese no-EK tendría una exactitud de 0,99). Sin embargo, el estudio está realizado con una población con una incidencia de EK superior a la de nuestro medio y en centros terciarios que podrían tener una prevalencia aún mayor que la población general de niños con fiebre de la que procede. En poblaciones con menor incidencia, el rendimiento sería menor. Otro factor que podría influir es el momento de realizar la analítica, que puede ser más precoz en nuevos pacientes de otras poblaciones, lo que también podría influir en el rendimiento del modelo.

Estos resultados van en la misma dirección que la de otro estudio realizado previamente², en el que se observó una pérdida del rendimiento tras la validación externa.

Aplicabilidad en la práctica clínica: aunque el modelo parece una herramienta útil para identificar de forma sencilla los pacientes con riesgo de padecer EK, existen dos inconvenientes que dificultan su aplicabilidad en nuestro medio.

En primer lugar, la población es muy diferente a la nuestra en términos de raza y factores ambientales, lo que podría sesgar el funcionamiento del modelo, basado en variables clínicas. En segundo, es necesario realizar una validación externa del modelo, especialmente en poblaciones más parecidas a la nuestra, antes de poder recomendar su uso, sobre todo teniendo en cuenta que, en condiciones de práctica habitual, algunos hechos como que los pacientes puedan ser evaluados de forma más precoz podrían mermar el rendimiento del modelo.

Conflicto de intereses de los autores del comentario: no existe.

BIBLIOGRAFÍA:

1. McCrindle BW, Rowley AH, Newburger JW, Burns JC, Bolger AF, Gewitz M, et al. Diagnosis, treatment, and long-term management of Kawasaki disease: a scientific statement for health professionals from the American Heart Association. *Circulation*. 2017;135:e927-99.
2. Hao S, Jin B, Tan Z, Li Z, Ji J, Hu G, et al.; Pediatric Emergency Medicine Kawasaki Disease Research Group. A classification tool for differentiation of Kawasaki disease from other febrile illnesses. *J Pediatr*. 2016;176:114-20.
3. Calcupedev. Herramienta de cálculo epidemiológico en pediatría. E. Ortega Páez. Comité de Pediatría Basada en la Evidencia de la AEP. 2019 [en línea] [consultado el 08/01/2023]. Disponible en www.aepap.org/calculadora-estudios-pbe/##

* Datos calculados por los revisores a partir de los datos originales con Calcupedev³.