

Inteligencia Artificial en IRA: la brecha entre la investigación y el mundo real

Dra. Daniela Ponce

*Professora Titular da Disciplina de Nefrologia do Faculdade de Medicina de Botucatu
Coordenadora do PPG Fisiopatologia em Clínica Médica - UNESP - Botucatu
Responsável Médica pela Unidade de Diálise do HC-FMB*

La injuria renal aguda (IRA) afecta alrededor de 200 pacientes por millón de habitantes, el 22% de los pacientes hospitalizados y más del 50% de los pacientes en la unidad de cuidados intensivos (UCI) [1]. El desarrollo de IRA se asocia con estancias hospitalarias prolongadas y un aumento del gasto relacionado con la salud [2]. Además, los supervivientes de IRA tienen un mayor riesgo de rehospitalización a los 30 días [2]. El uso de big data e inteligencia artificial (IA) puede ayudar a los trabajadores sanitarios a mejorar las decisiones en el tratamiento de pacientes con lesión renal aguda, reduciendo resultados indeseables [3].

El manejo de la IRA se refiere a las estrategias y prácticas utilizadas para mejorar la detección, prevención y tratamiento de la IRA en entornos clínicos [1]. Las prácticas de gestión también podrían incluir el uso de modelos y algoritmos predictivos. Este enfoque tiene el potencial de identificar pacientes de alto riesgo, implementar protocolos e intervenciones para controlar la IRA, utilizar sistemas de registros médicos electrónicos (EHR), un formato de datos digitales sobre la salud de los pacientes, para monitorear los datos relacionados con la IRA y educar a los proveedores médicos sobre las mejores prácticas para el tratamiento de la IRA.

Se han desarrollado y probado varias puntuaciones y modelos de predicción, que van desde la identificación del riesgo de IRA o resultados específicos, como el riesgo de nuevas puntuaciones de mortalidad por diálisis, hasta la identificación del riesgo de mortalidad [3, 4].

La mayoría de los modelos de predicción y el uso recurrente de la IA se desarrollan en entornos de unidades de cuidados intensivos (UCI) y en países desarrollados. Sin embargo, la mayoría están dirigidas a la mortalidad. La mayoría de los modelos de predicción centrados en el manejo médico son aún escasos. Algunas publicaciones aplicaron la IA en escenarios del mundo real, y aún menos estudios han evaluado los efectos de la IA en los resultados clínicos y organizacionales [3-5].

IA en la predicción del riesgo de lesión renal aguda

Estudios recientes han demostrado que la IRA puede ser predecible y prevenible si se identifican los factores de riesgo tempranos [1-3]. Aunque los factores de riesgo de IRA son complejos y a menudo se superponen, identificarlos con precisión brinda la oportunidad de prevenir incidentes y reducir los resultados negativos asociados con la IRA mediante el análisis de su correlación e impacto a través de modelos predictivos.

Los modelos de predicción de riesgos (RPM) son herramientas matemáticas que utilizan una combinación de variables para predecir el riesgo de un evento adverso basándose en datos existentes [3,4]. Los RPM pueden ser tanto diagnósticos como pronósticos de IRA. Para desarrollar RPM precisos y confiables se utilizan comúnmente algoritmos poderosos y herramientas avanzadas, como la Inteligencia Artificial (IA) y su subconjunto, el aprendizaje automático (ML). El desarrollo de RPM de AKI es relevante para los pacientes y sus familias, los proveedores de atención médica, los formuladores de políticas y los investigadores médicos. Estos modelos son esenciales para predecir el resultado de los pacientes con IRA y, a menudo, se utilizan para la estratificación del riesgo.

Los modelos precisos pueden estratificar eficazmente a los pacientes para identificar aquellos con alto riesgo de desarrollar IRA y así brindar atención especializada para prevenir una lesión renal adicional. La implementación clínica de los RPM se centra en predecir la incidencia de IRA con suficiente antelación para que las intervenciones relevantes mejoren los resultados de los pacientes. Se han dedicado muchas investigaciones a predecir la IRA, donde se consideran estrategias de prevención para quienes tienen un riesgo elevado. Los avances del ML han resultado en mejoras significativas en los RPM utilizados para estimar las incidencias de IRA. Además, la mayor disponibilidad de EHR hizo posible desarrollar RPM para estimar los factores de riesgo de IRA.

Estudios recientes han demostrado que no existe una combinación estándar de factores de riesgo ni una comprensión clara de su impacto y asociación con los episodios de IRA. Otros estudios revelaron que los RPM existentes son de poca utilidad debido a la diferencia en los factores de riesgo independientes utilizados para construir los modelos. Sin embargo, los proveedores médicos afirman que la IRA tiende a ocurrir en personas con factores de riesgo comunes y ciertas condiciones y han investigado exhaustivamente estos factores. Además, la introducción de EHR significa un gran paso adelante en la predicción de los factores que facilitan el riesgo de IRA. Un estudio reciente demostró el papel fundamental de los EHR en el descubrimiento de las interacciones entre los factores de riesgo, la reducción de los errores médicos y la mejora de los resultados de los pacientes [5].

Recientemente, Al-Absi et al analizaron 27 estudios publicados entre 2012 y 2023 y mostraron una heterogeneidad significativa en el diseño de los modelos, con variaciones en los entornos clínicos, las características de los pacientes, las regiones de la cohorte y los métodos estadísticos [3]. La mayoría de los modelos se desarrollaron para la IRA en pacientes hospitalizados, en particular aquellos sometidos a cirugía o en unidades de cuidados intensivos.

Los modelos de predicción de IRA generalmente se desarrollan utilizando datos retrospectivos de individuos con riesgo de IRA, incorporando información conocida de la función renal inicial y de seguimiento. Para todos los RPM analizados, los predictores identificados con mayor frecuencia fueron variables renales como SCr basal, bicarbonato, TFG, BUN y UO estimados, así como variables no renales como edad, sexo, índice de masa corporal, diabetes, insuficiencia hepática, ventilación mecánica, uso de medicamentos, insuficiencia cardíaca, uso de vasopresores, diuréticos, hemoglobina, recuento de glóbulos blancos, presión arterial, hipertensión e hipotensión (tabla 1).

La validación interna se realizó dividiendo aleatoriamente el conjunto de datos, pero este enfoque es limitado ya que los conjuntos de datos de derivación y validación comparten similitudes sustanciales. La validación externa rara vez se realizó en los estudios revisados y, cuando se realizó, reveló una menor precisión del modelo, lo que limitó la generalización de estos modelos en diversos centros y entornos médicos. La falta de validación externa podría deberse a desafíos para acceder a conjuntos de datos externos, incluidas preocupaciones sobre la privacidad del paciente, restricciones para compartir datos y problemas logísticos. Las diferencias en los métodos de recopilación de datos, las poblaciones de pacientes y los entornos sanitarios también complican la validación externa sin ajustes significativos al modelo original.

Estos datos enfatizan la necesidad de realizar más investigaciones para desarrollar modelos de predicción de riesgo confiables y precisos para la IRA que puedan aplicarse a la población general. Los datos para los modelos de predicción deben ser representativos. Debe provenir de diferentes poblaciones, etnias, condiciones de salud y de diferentes lugares.

IA en la predicción pronóstica de la IRA

El desarrollo de modelos de predicción que predigan con precisión a los pacientes con IRA con un alto riesgo de resultados adversos puede mejorar significativamente el manejo y el tratamiento de esos pacientes. Se han desarrollado varios modelos para la predicción del pronóstico de IRA utilizando métodos de ML, y sus rendimientos se compararon con los de los enfoques estadísticos

tradicionales, Proporcionar a los médicos información útil sobre otros eventos adversos o pronóstico a largo plazo.

Recientemente, Lin et al [4] publicaron una revisión sistemática y un metanálisis sobre la predicción del pronóstico de lesión renal aguda mediante métodos de aprendizaje automático. De los 27 estudios seleccionados, 21 compararon métodos tradicionales (como los sistemas de puntuación de gravedad Simplified Acute Physiology Score-APACHE II o la regresión logística) con modelos de aprendizaje automático. Excepto dos, los estudios de 21 estudios encontraron que los modelos ML tenían un AUROC o estadística C más alto que los modelos estadísticos tradicionales o puntuaciones de evaluación de la gravedad. Para la predicción de la mortalidad ($n = 20$), el AUROC de los modelos de aprendizaje automático en diferentes estudios osciló entre 0,679 y 0,924, y 18 de ellos superaron un AUROC $>0,750$. Los modelos de predicción de recuperación de la función renal o insuficiencia renal fueron menos precisos que los modelos de predicción de mortalidad, con un AUROC que oscilaba entre 0,640 y 0,846. Sólo 4 estudios proporcionaron datos suficientes para el metaanálisis de la recuperación funcional del riñón. Aunque el aprendizaje automático mostró una tendencia hacia una mayor discriminación de modelos, la diferencia no fue estadísticamente significativa (AUROC, 0,781; IC del 95 %, 0,674-0,861 vs. 0,734; IC del 95 %, 0,547-0,863) para diferentes modelos, lo que podría deberse al número limitado de estudios que predijeron la recuperación funcional del riñón.

Entre los estudios para la predicción de la mortalidad, 18 de ellos informaron características importantes para la mortalidad, incluida la edad, el nivel de creatinina sérica, el nivel de nitrógeno ureico sérico, la brecha aniónica y el recuento de glóbulos blancos, que se superponían como características predictivas de mortalidad que aparecieron en más de 6 estudios. Un conjunto clave de características predictivas de la mortalidad hospitalaria que incluye edad, nivel de creatinina sérica, nivel de nitrógeno ureico sérico, brecha aniónica y recuento de glóbulos blancos, estadio de IRA, tasa de filtración glomerular estimada y afecciones comórbidas como enfermedades cardiovasculares, cáncer y necesidad de ventilación mecánica. desempeñando un papel importante en la recuperación de la función renal o en la predicción de la insuficiencia renal. Las diversas etapas de la IRA, en particular la IRA que requiere diálisis, que es la forma más grave de IRA, se asociaron con una mayor mortalidad y falta de recuperación renal en los pacientes después de la cirugía. En el caso de la enfermedad renal aguda, factores como el nivel de creatinina sérica, las condiciones comórbidas y los medicamentos

desempeñaron un papel importante en la determinación de la persistencia de la enfermedad renal aguda (tabla 2).

En esta revisión sistemática integral de 27 estudios sobre el pronóstico de IRA, los autores encontraron que los modelos ML brindan una mejor discriminación en la predicción de la mortalidad que la recuperación de la función renal, y superan a los enfoques tradicionales en la predicción de la mortalidad entre pacientes con IRA

Los modelos de aprendizaje profundo que pueden identificar dinámicamente el riesgo de eventos adversos con características dependientes del tiempo incorporadas tienen un rendimiento de modelo superior a los modelos estáticos. Con la capacidad de identificar eventos adversos de alto riesgo en pacientes con IRA, los modelos de aprendizaje automático podrían ayudar a los médicos a realizar intervenciones oportunas y óptimas y, en última instancia, mejorar los resultados clínicos de los pacientes de alto riesgo. Dado su efecto potencial sobre la supervivencia del paciente, el pronóstico de los pacientes con IRA es un tema muy preocupante en el tratamiento posterior a la IRA.

No hubo estudios que involucren la IA en el momento del inicio de la KRT, por lo que sigue siendo controvertido en pacientes con IRA sin sobrecarga intensa de líquidos o trastornos metabólicos importantes, como acidosis, hiperpotasemia y uremia. El metanálisis no mostró diferencias en la tasa de mortalidad y la dependencia de la diálisis entre el inicio temprano o tardío de la KRT [1,2].

Los modelos de aprendizaje automático que predicen con precisión los resultados clínicos de los pacientes con IRA son capaces de (1) ayudar a los médicos a identificar tempranamente a los pacientes con alto riesgo de muerte o falta de recuperación y obtener la oportunidad de proporcionar intervenciones oportunas; (2) facilitar ensayos clínicos para pacientes con IRA con alto riesgo de resultados adversos, lo que podría reducir la mortalidad o la pérdida de la función renal; y (3) brindar asesoramiento médico e informar a los pacientes sobre los riesgos del deterioro de las condiciones. Además de los métodos tradicionales de aprendizaje automático, la tendencia creciente de utilizar modelos de aprendizaje profundo para identificar el riesgo de IRA es una estrategia popular para la detección temprana de pacientes con IRA. Los modelos de aprendizaje profundo pueden capturar información útil a partir de datos de series temporales, lo que demuestra un mejor rendimiento de predicción que los enfoques estadísticos tradicionales, los modelos de aprendizaje automático e incluso los expertos médicos.

GAP con implicaciones prácticas para los médicos

Si bien el desarrollo de RPM que utilizan ML ha demostrado un potencial significativo en la identificación de riesgos de IRA, la implementación de estos modelos en la práctica clínica presenta varios desafíos para los médicos. Un desafío clave es la complejidad de ciertos modelos, lo que los dificulta. Los médicos a menudo no están familiarizados con cómo los algoritmos de ML llegan a las predicciones, lo que puede generar aversión a la adopción de estas herramientas sin una comprensión clara de su funcionalidad. Esto resalta la importancia de las herramientas de IA explicables (XAI) que pueden proporcionar un razonamiento transparente para las predicciones y ayudar a mejorar la confianza y la usabilidad entre los médicos. Al ofrecer información sobre por qué ciertas características contribuyen a los resultados previstos, XAI puede ayudar a cerrar la brecha entre la IA y la toma de decisiones clínicas, haciendo que los modelos sean más aceptables para su aplicación en el mundo real.

Otro aspecto importante para la adopción clínica práctica es el equilibrio entre la complejidad y la usabilidad del modelo. Varios de los estudios revisados utilizaron modelos reducidos que priorizaban menos predictores pero clínicamente relevantes, manteniendo el poder predictivo y simplificando la implementación práctica. Los modelos reducidos son beneficiosos porque mantienen niveles de precisión similares a los modelos completos pero requieren menos variables clínicas, lo que los hace más fáciles de aplicar en un entorno sanitario del mundo real. Este enfoque puede ahorrar tiempo valioso, reducir los requisitos computacionales y mejorar la experiencia del médico al aplicar el modelo en un entorno clínico ajetreado.

Por lo tanto, los esfuerzos futuros deberían centrarse en el desarrollo de RPM que prioricen la relevancia clínica y la simplicidad para maximizar su usabilidad práctica sin sacrificar el rendimiento predictivo. La Figura 1 muestra las lagunas y las posibles soluciones para los modelos de predicción de AKI.

Limitaciones y conclusión

La aplicación de modelos de aprendizaje profundo en el riesgo y el pronóstico de la IRA sugiere que pueden predecir eficazmente la IRA y sus resultados adversos con cambios de tiempo, lo que muestra el potencial de ayudar a los médicos a identificar pacientes de alto riesgo. Sin embargo, la amplia gama de tipos de pacientes y entornos sanitarios de diferentes países dificulta establecer un enfoque estándar para el tratamiento de estas personas de alto riesgo. Para abordar estos desafíos, es esencial realizar más investigaciones que se centren en la validación externa, datos prospectivos, metodologías estandarizadas y modelos más simples y clínicamente factibles.

Desarrollar modelos confiables y prácticos requiere esfuerzos para incluir detalles clínicos que los registros médicos actuales pueden no capturar fácilmente. En particular, estos modelos deben validarse en diversos conjuntos de datos y hacerse de acceso público para maximizar su beneficio para los pacientes con IRA. Por lo tanto, los esfuerzos para desarrollar modelos de predicción de riesgo y pronóstico de IRA mejorados y ampliamente aplicables deben seguir siendo un foco clave de la investigación médica.

Referencias:

1. Susantitaphong P, Cruz DN, Cerda J, Abulfaraj M, Alqahtani F, Koulouridis I, et al. World incidence of AKI: a meta-analysis. *Clin J Am Soc Nephrol*. 2013;8(9):1482–93.
2. Chawla LS, Bellomo R, Bihorac A, Goldstein SL, Siew ED, Bagshaw SM, et al. Acute kidney disease and renal recovery: consensus report of the Acute Disease Quality Initiative (ADQI) 16 workgroup. *Nat Rev Nephrol*. 2017;13(4):241–57.
3. Al-Absi et al. Exploring the role of artificial intelligence in acute kidney injury management: a comprehensive review and future research agenda. *BMC Medical Informatics and Decision Making* (2024) 24:337
4. Lin Y, Shi S, Kong G. Acute Kidney Injury Prognosis Prediction Using Machine Learning Methods: A Systematic Review. *Kidney Med*. 7(1):100936. Published online November 15, 2024. doi: 10.1016/j.xkme.2024.100936
5. Holmes J, Allen N, Roberts G, Geen J, Williams JD, Phillips AO et al. Acute kidney injury electronic alerts in primary care - findings from a large population cohort. *QJM: An International Journal of Medicine* [Internet]. 2017 Sep 1 [cited 2023 Feb 14];110(9):577–82. <https://doi.org/10.1093/qjmed/hcx080>

Tabla 1. Conjunto clave de características predictivas para la IRA identificadas en modelos de predicción que utilizan IA

Género
Edad
Creatinina sérica basal
Urea sérica
Índice de masa corporal
Glóbulos blancos
bicarbonato
diuresis
hipotensión
Uso de vasopresores, diuréticos, antibióticos
Condiciones comórbidas: diabetes, insuficiencia hepática, ventilación mecánica, insuficiencia cardíaca, hipertensión

Tabla 2. Conjunto clave de características predictivas para la mortalidad en el hospital y la recuperación de la función renal identificadas en modelos de predicción que utilizan IA

Edad
Creatinina sérica basal
Urea sérica
Anion gap
Glóbulos blancos
Estadio de la IRA
Condiciones comórbidas: enfermedad cardiovascular, cáncer y necesidad de ventilación mecánica

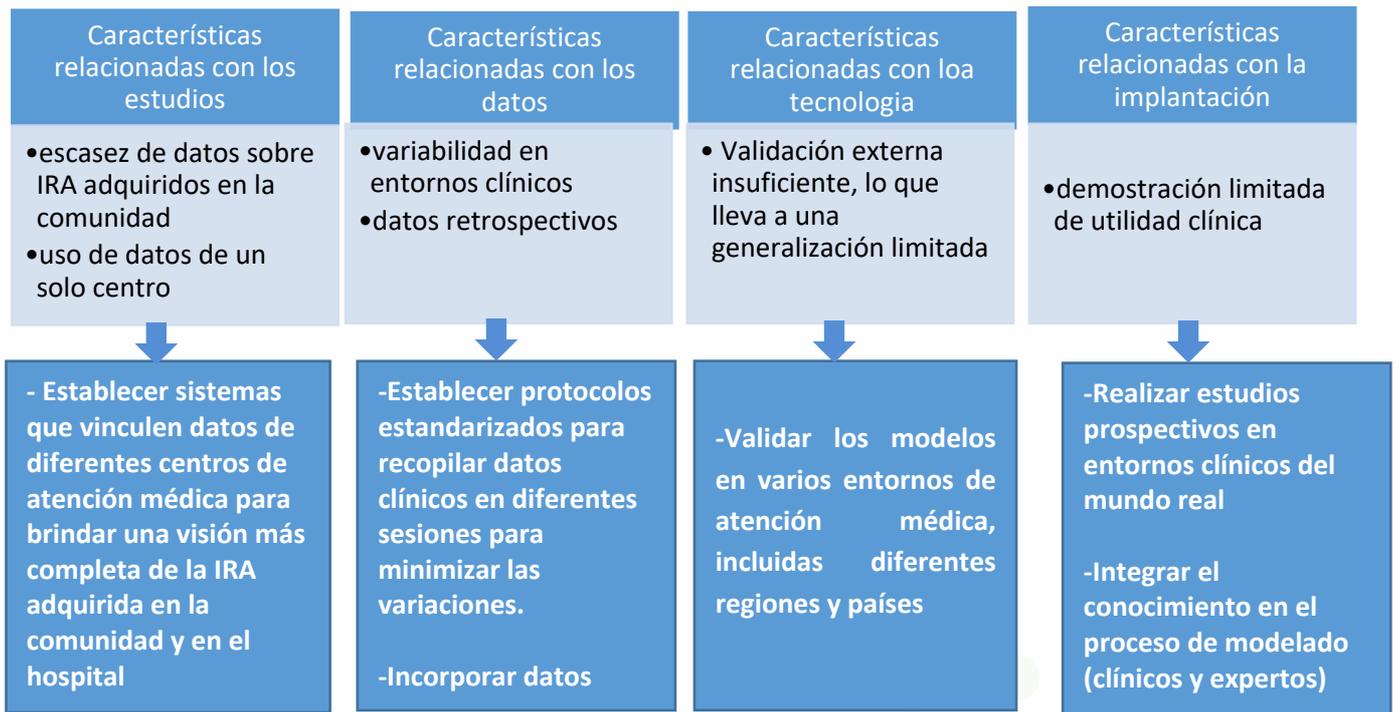


Figura 1. Gaps y potenciales soluciones para los modelos de predicción para la IRA

Inteligencia Artificial en IRA: La brecha entre la investigación y el mundo



Métodos	Hallazgos	Hallazgos
<p>Revisión narrativa de modelos de IRA aplicados en IA. (2012–2023)</p> <p>Se analizaron modelos de predicción y pronóstico de IRA</p> <p>Heterogeneidad entre estudios</p>	<p>PREDICCIÓN DE RIESGO</p> <p>Variables clínicas Edad, Cr basal, TFG, urea, BUN, diuresis, comorbilidades, medicamentos</p> <p>Uso de registros electrónicos y aprendizaje automático</p> <p>Limitada validación externa = menor aplicabilidad global</p> <p>Necesidad de modelos simples y comprensibles</p>	<p>PREDICCIÓN DE PRONÓSTICO</p> <p>Predicción de mortalidad AUROC 0.679–0.924 (ML > modelos clásicos)</p> <p>Recuperación renal Menor precisión en recuperación vs. mortalidad</p> <p>Modelos dinámicos Superiores a modelos estáticos</p> <p>Falta evidencia en momento óptimo para TRR</p>
<p>Intervención</p> <p>Algoritmos de aprendizaje automático</p> <p>Métodos estadísticos tradicionales</p> <p>Aplicación clínica limitada</p> <p>Escasa validación externa y evidencia en escenarios reales.</p>		