



Apuntes para la lectura crítica de un artículo sobre pronóstico

V. Barrio* y F. García López**

Grupo de Nefrología Basada en la Evidencia de la Sociedad Española de Nefrología. *Unidad de Nefrología. Fundación Hospital de Alarcón. **Unidad de Epidemiología Clínica. Hospital Universitario Puerta de Hierro. Madrid.

INTRODUCCIÓN

En el presente artículo se describen de forma esquemática los puntos más importantes a considerar para la lectura crítica de un estudio sobre pronóstico. Los estudios pronósticos abordan cuestiones del tipo de cuál es el riesgo de desarrollo de una insuficiencia renal terminal en un paciente con un filtrado glomerular estimado de 40 ml/min o si merece la pena efectuar una detección precoz de la proteinuria en la población general para reducir el riesgo futuro de desarrollo de insuficiencia renal avanzada. Estas preguntas contienen un elemento cualitativo (¿cuáles pueden ser los resultados?) y una escala temporal (¿cuándo podrán ocurrir?). Los puntos que se exponen a continuación son un desarrollo de las conocidas guías de lectura crítica de la literatura del grupo de la Medicina Basada en la Evidencia. Son tres los aspectos fundamentales para evaluar un estudio sobre pronóstico, en primer lugar su validez, en segundo lugar la importancia de los resultados y, por último, su relevancia y potencial aplicación a nuestro medio y a cada caso particular.

Antes de iniciar la exposición, y para centrar el tema, se discutirá brevemente qué tipo de estudios proporciona la mejor información sobre el pronóstico de un grupo de individuos o población con un problema de salud determinado o un factor de riesgo. Una revisión sistemática que combinara los resultados de todos los estudios sobre pronóstico de una determinada condición sería sin duda la mejor fuente de información. Sin embargo este tipo de artículos de revisión son escasos en la literatura y para evaluar su validez se debe recurrir a las normas de lectura crítica de las revisiones sistemáticas y técnicas de meta-análisis que ya se comentaron en otra revisión. Por lo tanto, y centrándonos en los estudios individuales, se debe mencionar en primer lugar los estudios de cohortes que consisten en la selección de una población con cier-

tas características con un seguimiento en el tiempo y monitorización del evento de interés, y que representan el paradigma de los estudios de pronóstico. Por citar un ejemplo conocido de todos recordemos la cohorte de Framingham y su importancia en la descripción de los factores de riesgo cardiovascular, así como de la importancia de la hipertensión como factor de riesgo de la insuficiencia renal. Los ensayos clínicos, en particular si el grupo control no recibe tratamiento activo, también pueden proporcionar información sobre el pronóstico de determinada condición, aunque en estos casos los criterios de inclusión suelen ser restrictivos, lo que reduce la validez externa de los resultados. Los estudios de casos y controles en los que se mide la exposición a un factor de forma retrospectiva, una vez ocurrido el evento de interés, son especialmente útiles en los casos en los que el evento ocurre de forma infrecuente o requiere un largo tiempo de seguimiento. Sin embargo la fuerza de la evidencia derivada de este tipo de estudios se encuentra limitada por los potenciales sesgos de selección y de medida inherentes a este tipo de diseño. En muchas ocasiones, se combina de forma cuantitativa la información sobre diagnóstico o pronóstico derivada de diferentes parámetros clínicos, de laboratorio o de imagen en los llamados modelos de predicción clínica que tampoco se analizan en esta revisión.

A. ¿SON VÁLIDOS LOS RESULTADOS DEL ESTUDIO PRONÓSTICO?

A.1. Preguntas de «eliminación»

Estas primeras preguntas sobre la validez interna del estudio considerado son de eliminación, ya que si la respuesta fuera negativa, no merecería la pena continuar con la evaluación del artículo.

1) ¿Se reunió una muestra adecuada de pacientes?

1.1. Definición del problema. En primer lugar, debe quedar claramente establecido y deli-

Correspondencia: Dr. Vicente Barrio Lucía
Fundación Hospital de Alarcón
C/ Budapest,1
28922 Alarcón (Madrid)
E-mail: vbarriol@senefro.org

mitado el problema objeto de estudio, se deben definir de forma explícita y estandarizada los criterios de diagnóstico de la patología o la enfermedad objeto del estudio pronóstico.

- 1.2. Diseño del estudio pronóstico. Se debe comprobar que el tipo de diseño es el más adecuado para responder a la pregunta planteada. Como se indicó arriba, los estudios de cohorte y los ensayos clínicos son los diseños más adecuados.
- 1.3. Selección de la muestra. Se debe valorar si la selección de la muestra de la población con la condición particular que se estudia es representativa, para no incurrir en sesgos de selección, lo que invalidaría, o al menos limitaría, la validez de los resultados. Esto significa que la muestra debería incluir todo el espectro de gravedad de la enfermedad en un momento dado, con objeto de que la muestra sea lo más representativa del conjunto de pacientes que padecen la enfermedad. Además, los distintos criterios de gravedad deberían estar claramente diferenciados y definidos.
- 1.4. Momento del inicio del seguimiento. Se debe valorar si los miembros de la cohorte afectados se encuentran en el mismo tiempo de evolución del proceso: por ejemplo, los resultados pueden ser muy dispares si la selección incluye tanto a pacientes incidentes como a prevalentes para una determinada condición. Por ejemplo, si el seguimiento incluye a pacientes prevalentes, puede ocurrir que algunos pacientes hayan fallecido o que otros se hayan recuperado antes de poder ser incluidos en el estudio. Por tanto, se debe comprobar que se incluyeron los pacientes del estudio en un momento uniformemente precoz de la enfermedad o de la situación objeto del estudio, idealmente desde el principio. Ejemplos de ello sería la fecha del trasplante renal en pacientes trasplantados o la fecha de inicio de tratamiento renal sustitutivo en pacientes con insuficiencia renal terminal.

2) *¿Fue completo el seguimiento del paciente?*

- 2.1. Duración del seguimiento. Se debe comprobar que el tiempo de seguimiento es suficientemente largo como para que ocurra el evento de interés. En caso contrario pueden quedar inadvertidos los efectos de determinadas exposiciones o factores de riesgo. Por el contrario, es muy valiosa la información si tras un tiempo de seguimiento muy prolongado apenas se han producido eventos.
- 2.2. Exhaustividad del seguimiento. Se debe comprobar que las pérdidas de seguimiento son pequeñas, o que al menos no tienen relación con el evento de interés. Habitualmente se acepta como cifra óptima de pérdidas de seguimiento un nivel inferior al 5% mientras que cifras superiores al 20% se consideran inaceptables. El problema surge cuando el motivo del abandono del seguimiento tiene que ver con el pronóstico, por ejemplo si el paciente muere, en cuyo caso se infringirían las premisas que subyacen a todos los análisis estadísticos de supervivencia, con la consecuencia de resultados sesgados. Por tanto, si el número de pérdidas de seguimiento es importante, se puede indagar si existen diferencias en algún factor pronóstico importante entre los que abandonan o se pierden de seguimiento, aunque esto sólo será posible si está documentada esta característica al inicio del seguimiento. Otra forma de ajustar consiste en realizar un análisis de sensibilidad suponiendo los escenarios más desfavorables (con el supuesto del "peor de los casos") y viendo en qué sentido y grado estos supuestos modifican la estimación de los resultados. Por ejemplo, en estos análisis se toman los casos perdidos en el seguimiento como fallecidos y se comparan sus resultados con los del análisis principal: si discrepan se pone en entredicho el resultado principal; si concuerdan, lo refuerzan.

A.2. Preguntas de detalle

3) *¿Se evaluaron los desenlaces de modo objetivo y enmascarado?*

- 3.1. Evaluación objetiva. Pocas veces el diagnóstico del evento es tan claro como en el caso de la muerte. En la mayoría de las ocasiones (como las causas de muerte, el rechazo agudo del injerto, las complicaciones cardiovasculares) los diagnósticos están sometidos a sesgos de medida y de clasificación y son necesarios criterios diagnósticos específicos, definidos de antemano.
- 3.2. Evaluación enmascarada. La evaluación del desenlace, de manera ideal, la deberían practicar clínicos que fueran desconocedores de las características de los pacientes y de sus factores pronósticos posibles. Por ejemplo, un patólogo que leyera una biopsia de un injerto renal para buscar criterios de rechazo agudo en un protocolo de investigación pronóstica

debería estar enmascarado a los síntomas y determinaciones analíticas del paciente.

4) *¿Se tuvieron en cuenta en el análisis estadístico todos los factores pronósticos importantes?*

4.1. Ajuste estadístico. El inconveniente principal de los estudios de cohortes es que los distintos grupos pueden diferir en variables distintas a las que determinan las características pronósticas que se estudian, las que dividen a los grupos en grupo de exposición y grupo control. Por tanto, el análisis estadístico no puede limitarse a comparar la frecuencia o las tasas de casos que desarrollan el evento en el grupo exposición y en el grupo control sino que tiene que tener en cuenta la presencia de otras variables de confusión, asociadas a la presencia de exposición y al evento, que pueden alterar la asociación entre la exposición y el evento. En el análisis estadístico, se debe ajustar, mediante análisis estratificados, análisis multivariantes o técnicas especiales (como la probabilidad de similitud —*propensity score*— entre los grupos), los factores pronósticos o de confusión, aunque en muchas ocasiones estos no tengan relación casual con el evento de interés, para ver cómo se modifican los resultados globales.

4.2. Desarrollo y validación en los modelos de predicción clínica. Cuando el propósito del estudio pronóstico es el establecimiento de un modelo de predicción clínica, se pretende conocer qué factores se asocian con el pronóstico, al margen de su relación causal o no con el desenlace. En esos casos, el hallazgo de una asociación puede ser debido al juego del azar, de modo que ese hallazgo debería validarse en un grupo independiente de pacientes con la misma enfermedad objeto de estudio para confirmar el poder predictivo y la reproducibilidad de los resultados. La primera cohorte sería la población de desarrollo del modelo pronóstico y la segunda cohorte sería la población de validación.

Una vez comprobada la validez del estudio de pronóstico con el esquema propuesto, se puede continuar, evaluando su importancia y aplicación a cada caso particular.

B. ¿CUÁLES SON LOS RESULTADOS?

5. *¿Cuál es la probabilidad de los resultados a lo largo del tiempo?*

5.1. Frecuencia de ocurrencia del desenlace. Habitualmente, los resultados de un estudio de pronóstico se expresan como porcentaje de supervivencia a un tiempo dado, como mediana de supervivencia o bien como curva de supervivencia de Kaplan-Meier, está última siendo la más completa y aconsejable ya que muestran en curso evolutivo a lo largo del tiempo. Por ejemplo, esta información se puede aplicar para estimar la mediana de supervivencia de un paciente diabético anciano que inicia diálisis (el tiempo en que sobrevive el 50% de la muestra) o la probabilidad de supervivencia del injerto de un paciente de 60 años que recibe un trasplante renal.

5.2. Diagrama de flujo del estudio. Siempre es útil conocer el número de participantes que se incluyeron en el estudio, los que se mantuvieron hasta el final o hasta el desarrollo del desenlace, y los que abandonaron o se perdieron en el seguimiento, con sus motivos. La representación gráfica mediante un diagrama de flujo ilustra este proceso.

5.3. Tasas de incidencia. Otra forma alternativa de expresar los resultados es mediante una densidad de incidencia, es decir, mediante tasas por unidad de tiempo (por ejemplo, número de muertos por años-paciente, infecciones peritoneales por meses-paciente o número de hipotensiones intradiálisis por horas-paciente).

5.4. Medidas de efecto. La comparación entre los resultados entre el grupo con la variable pronóstica de interés y el grupo control se debe expresar mediante medidas de efecto. Las medidas de efecto más frecuentes son la razón de riesgos proporcionales (*hazard ratio*), la razón de riesgos, la razón de tasas, la diferencia de riesgos y la diferencia de tasas. Como los valores de probabilidad, valores de P, transmiten menos información que las medidas de efecto y son muy sensibles al tamaño de la muestra, idealmente no deben describirse cuando se comparan unos grupos con otros, o al menos no deben constituir la base de la comparación.

5.5. Análisis de sensibilidad. Como se indicó arriba, cuando el número de perdidos en el seguimiento es alto se suelen efectuar análisis de sensibilidad con «el peor de los casos» para evaluar la consistencia de los resultados principales.

5.6. Análisis de subgrupos. Cuando hay subgrupos que difieren en sus características pronósticas es importante conocer si la frecuencia de ocu-

rrencia del desenlace o sus medidas de efecto difieren entre ellos. Como en los análisis de subgrupos de otros diseños, estos análisis deberían programarse de antemano y no estar condicionados por los datos.

6) *¿Fueron precisos los resultados?*

6.1. Intervalos de confianza. Para expresar la incertidumbre del resultado, se emplea el intervalo de confianza al 95%, que representa el rango de valores en que se encontrarían los resultados si se efectuaron 100 estudios iguales pero en muestras diferentes. Cuanto más estrecho sea el intervalo de confianza, mayor será la precisión de la estimación. Señalar que en las curvas de supervivencia los periodos de seguimiento iniciales son más precisos al incluir un mayor número de pacientes en seguimiento. Las medidas de efecto deben acompañarse siempre de sus intervalos de confianza correspondientes.

C. ¿PUEDEN AYUDARNOS ESTOS RESULTADOS?

Finalmente, estaremos en condiciones de evaluar la relevancia y potencial aplicación a nuestro medio y a cada caso particular de los resultados obtenidos en el estudio de pronóstico de una enfermedad determinada.

En este apartado, son tres los aspectos a considerar.

7) *¿Se pueden aplicar estos resultados en nuestro medio?*

7.1. Tipo de pacientes y ámbito del estudio pronóstico.

Se debe comprobar si las características demográficas y clínicas de nuestro paciente particular coinciden con las de la población utilizada para establecer el pronóstico, es decir si nuestro paciente hubiera podido ser incluido en el estudio. Siempre existirán algunas diferencias, y lo que se debe calibrar es su importancia y potencial impacto en el resultado.

8) *¿Se tuvieron en cuenta todos los desenlaces de importancia clínica?*

8.1. Desenlaces de importancia clínica. Es importante evaluar si se analizan todos los desenlaces de importancia, no sólo aquellos más llamativos, sino también otros que afectan directamente al paciente, como su calidad

Tabla I. Guía para la lectura de un estudio pronóstico

-
- A) ¿Son válidos los resultados del estudio pronóstico?
 - A.1. Preguntas de «eliminación»
 - 1. ¿Se reunió una muestra adecuada de pacientes?
 - Definición del problema
 - Objetivo específico de la revisión
 - Diseño del estudio pronóstico
 - Selección de la muestra
 - Momento del inicio del seguimiento
 - 2. ¿Fue completo el seguimiento del paciente?
 - Duración del seguimiento
 - Exhaustividad del seguimiento
 - A.2. Preguntas de detalle:
 - 3. ¿Se evaluaron los desenlaces de modo objetivo y enmascarado?
 - Evaluación objetiva
 - Evaluación enmascarada
 - 4. ¿Se tuvieron en cuenta en el análisis estadístico todos los factores pronósticos importantes?
 - Ajuste estadístico
 - Desarrollo y validación en los modelos de predicción clínica
 - B) ¿Cuáles son los resultados?
 - 5. ¿Cuál es la probabilidad de los resultados a lo largo del tiempo?
 - Frecuencia de ocurrencia del desenlace
 - Diagrama de flujo del estudio
 - Tasas de incidencia
 - Medidas de efecto
 - Análisis de sensibilidad
 - Análisis de subgrupos
 - 6. ¿Fueron precisos los resultados?
 - Intervalos de confianza
 - C) ¿Pueden ayudarnos estos resultados?
 - 7. ¿Se pueden aplicar estos resultados en nuestro medio?
 - Tipo de pacientes y ámbito de los estudios del estudio pronóstico
 - 8. ¿Se tuvieron en cuenta todos los desenlaces de importancia clínica?
 - información importante para la toma de decisiones médicas
 - información importante o irrelevante por la toma de decisiones médicas.
-

de vida relacionada con la salud u otros aspectos como la hospitalización o las cargas económicas.

9) *¿Son útiles los resultados para tranquilizar o aconsejar a los pacientes?*

9.1. Información importante para la toma de decisiones médicas. Finalmente, habría que considerar si esta nueva información sobre el pronóstico va a tener alguna relevancia sobre decisiones importantes en el manejo clínico del paciente, por ejemplo a la hora de indicar

monitorización de su estado y progresión con alguna prueba diagnóstica o bien en la decisión de iniciar o indicar un tratamiento.

- 9.2. Información pronóstica que no influye en la toma de decisiones médicas. Incluso si el conocimiento del pronóstico no es útil para tomar decisiones que alteren el curso de la evolución del paciente, sí puede serlo para tranquilizar o aconsejar al paciente o a su familia.

BIBLIOGRAFÍA

- Laupacis A, Wells G, Richardson WS, Tugwell P for the Evidence-Based Medicine Working Group. *JAMA* 272: 234-237, 1994.
- Jaeschke R, Guyatt G, Sackett DL: Users' guides to the medical literature. III. How to use an article about a diagnostic test. A. Are the results of the study valid? Evidence-Based Medicine Working Group. *JAMA* 271: 389-391, 1994.
- Guyatt GH, Sackett DL, Cook DJ: Users' guides to the medical literature. II. How to use an article about therapy or prevention. B. What were the results and will they help me in caring for my patients? Evidence-Based Medicine Working Group. *JAMA* 271: 59-63, 1994.
- McAlister FA, Straus SE, Guyatt GH, Haynes RB. Users' guides to the medical literature: XX. Integrating research evidence with the care of the individual patient. Evidence-Based Medicine Working Group. *JAMA* 283: 2829-36, 2000.
- García López F, Quereda Rodríguez-Navarro C: Apuntes para la lectura crítica de una revisión sistemática. *Nefrología* 27: 417-24, 2007.
- Sackett DL, Richardson WS, Rosenberg W, Haynes RB: Medicina basada en la evidencia: cómo ejercer y enseñar la MBE. Madrid: Churchill Livingstone, 1997.
- Guyatt G, Rennie D: Users' guides to the medical literature: essentials of evidence-based clinical practice. Chicago: AMA Press, 2002.
- Guyatt G, Rennie D: Users' guides to the medical literature: a manual for evidence-based clinical practice. Chicago: AMA Press, 2002.
- CASPe (Critical Appraisal Skills Program en España): programa de habilidades en lectura crítica, España. <http://www.redcaspe.org/herramientas/index.htm>
- Centre for Evidence-Based Medicine, Oxford-Centre for Evidence Based Medicine. The EBM toolbox. <http://www.cebm.net/toolbox.asp>
- User's guides interactive: an online tool to guide clinicians in the appraisal and application of evidence into their everyday practice. <http://www.usersguides.org/>