



Fundamentos de Medicina Basada en la Evidencia

Valoración crítica de estudios sobre reglas de predicción clínica

Carlos Ochoa Sangrador. Servicio de Pediatría. Hospital Virgen de la Concha. Zamora (España).
Correo electrónico: cochoas@meditex.es

Términos clave en inglés: logistic models; models, theoretical; decision support techniques

Términos clave en español: modelos logísticos; modelos teóricos; técnicas de apoyo a la decisión

Conflicto de intereses: no existe.

Fecha de recepción: 18 de mayo de 2009
Fecha de aceptación: 19 de mayo de 2009

Fecha de publicación: 1 de junio de 2009

Evid Pediatr. 2009; 5: 52 doi: vol5/2009_numero_1/2009_vol5_numero1.20.htm

Cómo citar este artículo

Ochoa Sangrador C. Valoración crítica de estudios sobre reglas de predicción clínica. Evid Pediatr. 2009; 5: 52.

Para recibir Evidencias en Pediatría en su correo electrónico debe darse de alta en nuestro boletín por medio del ETOC en <http://www.aepap.org/EvidPediatr/etoc.htm>

Este artículo está disponible en: http://www.aepap.org/EvidPediatr/numeros/vol5/2009_numero_2/2009_vol5_numero2.20.htm
EVIDENCIAS EN PEDIATRIA es la revista oficial del Grupo de Pediatría Basada en la Evidencia de la Asociación Española de Pediatría de Atención Primaria. © 2005-09. Todos los derechos reservados. ISSN : 1885-7388

Valoración crítica de estudios sobre reglas de predicción clínica

Carlos Ochoa Sangrador. Servicio de Pediatría. Hospital Virgen de la Concha. Zamora (España).
Correo electrónico: cochoas@meditex.es

Introducción

Podemos definir la regla de predicción clínica (RPC) como una herramienta compuesta por un conjunto de variables de la historia clínica, exploración física y pruebas complementarias básicas, que nos estima la probabilidad de un evento, nos sugiere un diagnóstico o nos predice una respuesta concreta a un tratamiento¹. Esta definición resulta igualmente aplicable para las denominadas "guías de predicción clínica" o "reglas de decisión clínica". La "predicción" implica ayudar al médico a tomar la mejor decisión sobre un acontecimiento clínico futuro. La "decisión" implica dirigir al médico para que realice un tipo de acción específica, diagnóstica o terapéutica. La aplicación de las reglas de predicción clínica a veces deriva en una decisión y a veces en una predicción, pero a menudo también en una probabilidad o un cociente de probabilidades que el médico aplica al problema diagnóstico en cuestión. En este último caso, el término de "regla de diagnóstico clínico" o "guía de diagnóstico clínico" sería más preciso².

En un artículo previo hemos revisado los fundamentos del diseño y evaluación de las RPC³. En él describimos cómo para diseñar una RPC deben seguirse varios pasos: 1) elegir el diagnóstico o evento a predecir y el escenario clínico o población en el que se quiere aplicar; 2) evaluar las variables predictivas en una muestra adecuada de pacientes, diseñando una RPC que recoja y pondere las variables finalmente válidas (derivación o diseño); 3) validar la RPC en una nueva muestra de pacientes similar a la de derivación, contrastando sus predicciones con el patrón de referencia que identifica el diagnóstico o evento de interés; y 4) aplicar la RPC en la práctica clínica examinando su rendimiento clínico.

Las RPC no pueden sustituir nuestra experiencia y conocimientos, sin embargo nos facilitan información estructurada sobre la capacidad predictiva de cada uno de los síntomas, signos y pruebas complementarias implicados en un determinado escenario clínico, y cómo combinarlos para cuantificar de forma precisa nuestras estimaciones diagnósticas y pronósticas.

En la práctica clínica, la información de la que disponemos para juzgar la validez y aplicabilidad de una RPC procederá de artículos científicos en los que se refleje su diseño o evaluación. Por ello, ante un escenario concreto en el que pensamos emplear una RPC tendremos que buscar estos artículos y, a través de su lectura crítica, juzgar la validez, importancia clínica y aplicabilidad de la RPC objeto de evaluación.

Búsqueda de estudios sobre RPC

Para buscar información sobre una RPC podemos explorar distintas bases de datos, tanto de publicaciones primarias como secundarias. En la búsqueda usaremos descriptores

relacionados con la predicción clínica combinados con los de la enfermedad o evento que se trata de predecir. Lamentablemente, no existe un descriptor específico en el repertorio de descriptores (MeSH) de la base de datos MEDLINE de la *National Library of Medicine*. Se han usado diversas estrategias de búsqueda, en las que el descriptor más específico parece ser "Decision Support Techniques" y el término aislado más sensible "predict\$"⁴. En la tabla 1 se detallan una serie de alternativas útiles.

Podemos encontrar artículos en los que se haya realizado el diseño, validación y análisis del impacto clínico de una RPC, sin embargo lo más frecuente es encontrar por separado estos estudios o a lo sumo estudios conjuntos de diseño y validación limitada.

Valoración crítica de estudios sobre RPC

Una vez encontrado el artículo o artículos debemos realizar una valoración de los mismos. Esta valoración va a resultar fundamental ya que a menudo la calidad de los trabajos publicados es deficiente, no se ajustan al escenario clínico que se trata de resolver, tienen errores metodológicos que comprometen los resultados o éstos son presentados de forma que limitan su correcta interpretación. Los fundamentos metodológicos de estos estudios han sido analizados en un artículo previo, en el que se justifican los criterios que a continuación repasaremos.

Diversos textos han analizado los estándares que deben cumplir estos estudios^{1,2,5,6}. Seguiremos los pasos habitualmente propuestos por el *Evidence Based Medicine Working Group*(2): ¿son válidos los resultados del estudio?, ¿cómo son los resultados?, ¿son aplicables a tu medio?

¿Son válidos los resultados del estudio?

El primer aspecto que hay que establecer es si la RPC responde a una pregunta claramente definida, en términos de población, diagnóstico o evento a predecir, variables predictivas consideradas y criterio de valoración de los resultados. Si alguno de estos elementos no están definidos o no se ajustan al escenario clínico que queremos resolver, el artículo que estamos valorando puede que no nos sea útil.

Debemos examinar las variables con potencial capacidad predictiva consideradas en el modelo (también las que han sido excluidas). Deben ser clínicamente importantes, fiables (precisión, reproducibilidad, concordancia intra e interobservador), estar definidas claramente y por adelantado. Además, el número de variables incluidas en la RPC no debe ser excesivo: para validar una RPC es conveniente una muestra con un número de eventos que supere 10 veces el número de variables predictivas, lo que evita fenómenos de sobreajuste.

La muestra que se emplee para diseñar la RPC debe

ser representativa de la que será objeto de aplicación. Además, deben especificarse todas sus características, para facilitar la interpretación de la aplicabilidad de los resultados a otras poblaciones.

El resultado de interés de la predicción clínica debe ser válido: el diagnóstico o evento que trata de predecir la RPC debe ser relevante, estar claramente definido y ser medido con procedimientos adecuados. Si en un estudio el patrón de referencia clasifica incorrectamente o deja sin clasificar a un grupo de pacientes, los resultados no serán válidos.

Además, durante el proceso de desarrollo del modelo o su validación, las variables predictivas y el resultado que se trata de predecir deben ser recogidos de forma prospectiva e independiente (interesa que quien recoja las variables predictivas desconozca el resultado del patrón de referencia y viceversa).

Los métodos empleados en el diseño de la RPC deben ser explícitos y adecuados, detallando los procedimientos estadísticos empleados en el diseño de la RPC. Por último, para conocer la reproducibilidad y aplicabilidad de la RPC, los resultados deben ser validados en una población diferente de la utilizada para su diseño.

¿Cómo son los resultados?

Si el apartado anterior nos indica que el estudio es válido daremos un paso más, analizando los resultados. Comenzaremos fijándonos en el método empleado para categorizar a los pacientes (categorización). Los métodos más usados serán las escalas de puntuación, las fórmulas de regresión o los diagramas de flujo diseñados mediante partición reiterativa.

A continuación buscaremos los indicadores que describen la validez de la RPC. Los valores de sensibilidad, especificidad, valores predictivos, cocientes de probabilidades o áreas bajo la curva ROC nos informarán de cómo discrimina el diagnóstico o evento a predecir (discriminación). Si dichos estimadores no se han calculado, trataremos de calcularlos a partir de los resultados del estudio. Si bien es frecuente que los trabajos presenten las estimaciones de sensibilidad o especificidad, e incluso de los valores predictivos, pocos trabajos estiman los cocientes de probabilidades. El interés de estos estimadores radica en su utilidad para el cálculo de probabilidades ajustado a entornos con prevalencias distintas de las encontradas en los estudios de diseño o validación.

También buscaremos si en el estudio se ha calibrado la RPC, empleando alguna técnica que nos muestre cómo se ajustan el rango de predicciones del modelo a los datos reales. Habitualmente se representarán con métodos de bondad de ajuste, tablas o representaciones gráficas. Esto será especialmente importante cuando la predicción comprenda múltiples categorías o valores continuos.

Por último, considerando que todas las estimaciones se han hecho sobre muestras, tendremos que examinar la precisión de los resultados, si se presentan los intervalos de confianza, cómo es su amplitud, y, en los casos que no se presenten, si podemos calcularlos.

¿Son aplicables a tu medio?

Una vez evaluada la validez de los resultados y su magnitud, tenemos que juzgar su reproducibilidad, factibilidad y aplicabilidad. Si una RPC ha sido validada en varias muestras de pacientes de entornos clínicos diferentes con buenos resultados, la podremos considerar generalizable y reproducible. Para juzgar la aplicabilidad de la RPC tendremos en cuenta si los pacientes incluidos en las muestras de derivación y validación son similares a nuestros pacientes (mirar criterios de inclusión y exclusión). Asimismo, nos interesa valorar la factibilidad de la RPC. Algunos modelos predictivos requieren cálculos complejos, especialmente aquellos que usan ecuaciones de regresión, o información clínica que por limitaciones de medios o tiempo no podemos recoger en nuestra práctica clínica.

Por último, tendremos que valorar el impacto clínico de la RPC. Ya hemos comentado anteriormente que es excepcional encontrar trabajos en los que se evalúe el rendimiento clínico con variables de resultados como: duración de la enfermedad, incidencia de complicaciones, supervivencia, costes, etc. Por ello, el rendimiento clínico ha de ser muchas veces estimado de forma indirecta, considerando si los resultados de la RPC modificarán el manejo de nuestros pacientes, con cambios en la indicación de procedimientos diagnósticos o terapéuticos, si influirán en su pronóstico o evolución, si los errores de la predicción (falsos positivos o negativos) tienen consecuencias potencialmente graves, y si su uso afectará al consumo de recursos o costes. La ponderación global de costes y beneficios deberá hacerse tanto desde la perspectiva del médico, como del paciente, el sistema sanitario y la sociedad.

Existen diversas herramientas que facilitan los pasos anteriormente mencionados. En la tabla 2 se presenta una versión modificada de las preguntas elaboradas por el grupo CASPe como ayuda para la valoración crítica de estudios sobre RPC, inspiradas en la propuesta del *Evidence Based Medicine Working Group*².

Criterios de calidad. Niveles de evidencia.

Como resultado de la valoración crítica, si la RPC es válida, sus resultados tienen importancia clínica y son aplicables a nuestro medio, decidiremos su implantación en nuestra práctica clínica o su incorporación a guías de práctica clínica. Para este supuesto nos va a interesar disponer de algún sistema de jerarquización de la calidad de la evidencia que sustenta la RPC.

En la tabla 3 resumimos los requerimientos metodológicos de los estudios de evaluación de RPC, presentados como criterios de calidad. Estos criterios serían aplicables a la valoración de trabajos en los que se hace el diseño y validación de una RPC. Si el trabajo aborda exclusivamente una de las fases, diseño o validación, los criterios requieren una lectura adaptada. Para estudios en los que se evalúa el rendimiento o impacto clínico de la aplicación de la guía, los criterios de calidad dependerán del diseño elegido en el estudio (ej. ensayo clínico, estudio de co-

hortes, etc.).

Algunos autores han hecho revisiones metodológicas de estudios de diseño o validación de RPC, analizando su calidad metodológica y el grado de cumplimiento de los estándares propuestos^{1,5,7}. Las principales deficiencias encontradas son: la ausencia de enmascaramiento o independencia en la recogida de información, la no valoración de la precisión o reproducibilidad de las variables predictivas, el uso de muestras que no representan el espectro habitual de pacientes, la falta de estudios de validación y la no valoración del impacto clínico de la aplicación de las RPC. Otros problemas menos frecuentes son: limitaciones en la definición de las variables predictivas o del patrón de referencia, en el grado de descripción de las muestras empleadas o en los métodos estadísticos empleados.

En función del grado de desarrollo alcanzado por una RPC podremos calificar el nivel de evidencia que la apoya. En la tabla 4 se presenta el sistema de jerarquización de la evidencia, propuesto por el *Evidence-Based Medicine Working Group*². Sobre esta base, Seilly et al⁷ proponen un nivel superior suplementario en el que entrarían RPC que han mostrado su impacto clínico en estudios realizados en diferentes entornos, mejorando las decisiones médicas

en un amplio espectro de pacientes.

Conflicto de intereses de los autores del comentario:
no existe

Bibliografía:

- 1.- Laupacis A, Sekar N, Stiell IG. Clinical prediction rules: a review and suggested modifications of methodological standards. *JAMA*. 1997; 277: 488-94.
- 2.- McGinn TG, Guyatt GH, Wyer PC, Naylor CD, Stiell IG, Richardson WS. Users' guides to the medical literature, XXII: how to use articles about clinical decision rules: Evidence-Based Medicine Working Group. *JAMA*. 2000; 284: 79-84.
- 3.- Ochoa Sangrador C. Diseño y evaluación de Reglas de Predicción Clínica. *Evid Pediatr*. 2009;5:28.
- 4.- Ingui BJ, Rogers MA. Searching for clinical prediction rules in MEDLINE. *J Am Med Inform Assoc*. 2001; 8: 391-7.
- 5.- Wasson JH, Sox HC, Neff RK, Goldman L. Clinical prediction rules: application and methodological standards. *N Engl J Med*. 1985; 313: 793-9.
- 6.- Perrier A, Perneger T. Les règles de prédiction: un guide de lecture. *Rev Mal Respir*. 2004; 21: S40-6.
- 7.- Reilly BM, Evans AT. Translating clinical research into clinical practice: impact of using prediction rules to make decisions. *Ann Intern Med*. 2006; 144: 201-9.

Tabla 1.- Búsquedas bibliográficas de estudios sobre reglas de predicción clínica

<p>Ejemplo de búsqueda:</p> <ol style="list-style-type: none"> 1.- Descriptores relacionados con la predicción clínica: (((“Logistic Models”[Mesh] OR “Models, Theoretical”[Mesh]) OR “Probability”[Mesh]) OR “Multivariate Analysis”[Mesh]) OR “Decision Support Techniques”[Mesh] 2.- Descriptores relacionados con la enfermedad o evento de interés; ejemplo (“Pneumonia, Viral”[Mesh]) 3.- Combinar “1” AND “2” 4.- Si aparecen demasiadas citas combinar la búsqueda anterior con el descriptor “Sensitivity and Specificity”[Mesh] (“3” AND “4”)
<p>Descriptores relacionados con las RPC empleados en la opción “Clinical queries” de PubMed:</p> <ul style="list-style-type: none"> - Búsqueda específica: (validation[tiab] OR validate[tiab]) - Búsqueda sensible: (predict*[tiab] OR predictive value of tests[mh] OR scor*[tiab] OR observ*[tiab] OR observer variation[mh])
<p>Otros descriptores aplicables a búsquedas en texto libre:</p> <ul style="list-style-type: none"> - “clinical prediction rules”; “rules”; “rule*”; “predict*”; “decisión rules”; “clinical*”; “outcome*”; “risk*”; “scor*”

Tabla 2.- Preguntas y pistas para entender las reglas de predicción clínica (modificado de <http://www.redcaspe.org/herramientas/lectura/11RPC-PRESAG.pdf>)

A ¿Son válidos los resultados del estudio?

- 1 ¿La regla responde a una pregunta bien definida?
 - ¿Se define claramente el tipo de Pacientes a los que se le aplica la regla?
 - ¿Están adecuadamente descritas las Variables predictoras?
 - ¿El desenlace (Outcome) es relevante y tiene sentido clínico? (El desenlace se puede expresar como una probabilidad o un curso de acción)
- 2 ¿La población a estudio de la que se derivó la regla, incluyó un espectro adecuado de pacientes?
 - ¿Es adecuado el método de selección de pacientes?
 - ¿Está adecuadamente representado el espectro de pacientes en los que tiene sentido aplicar la regla?
- 3 ¿Se validó la regla en un grupo diferente de pacientes?
 - No basta con que la regla "funcione" en la población a partir de la cual se ha derivado.
 - La validación se realizó con pacientes similares o distintos a aquellos con los que se generó.
- 4 ¿Hubo una evaluación ciega del desenlace y de las variables predictoras?
 - ¿Las personas que valoraban el resultado conocían los datos clínicos?
 - ¿Las personas que medían las variables predictoras conocían el desenlace?
- 5 ¿Se midieron las variables predictoras y el desenlace en todos los pacientes?
 - ¿Están bien descritas las exclusiones?
 - A veces el desenlace no se puede medir de la misma forma en todos los pacientes.
- 6 ¿Se describen los métodos de derivación y validación de la regla?
 - ¿Se incluyen las variables importantes y los criterios de positividad?
 - ¿Se describe la reproducibilidad de las medidas?

B ¿Cuáles son los resultados?

- 7 ¿Se puede calcular el rendimiento de la RPC?
 - Los resultados pueden presentarse como: S, Esp, CP+, CP-, curva ROC, curvas de calibración, tasas esperadas/observadas, etc.
- 8 ¿Cuál es la precisión de los resultados?
 - Es el momento para reflexionar sobre el tamaño de la muestra y el número de variables de la RPC.
 - ¿Es suficientemente robusta la regla? (se ha intentado refinar: análisis de sensibilidad, etc.)

C ¿Son los resultados aplicables al escenario?

- 9 ¿Serán satisfactorios en el ámbito del escenario la reproducibilidad de la RPC y su interpretación?
 - Considera si el ámbito de estudio es demasiado diferente al del escenario.
- 10 ¿Es aceptable la regla en este caso?
 - Considera la facilidad de uso, la disponibilidad de la RPC y los costes.
 - Considera si la RPC tiene sentido clínico.
- 11 ¿Modificarán los resultados la conducta clínica, los resultados en salud o los costes?
 - Desde la perspectiva del escenario, si la RPC no va a cambiar la actitud la RPC es (al menos) inútil.
 - Al margen de tu opinión puede haber estudios que exploren el impacto de la RPC (sobre costes o sobre resultados de salud).
 - Considera cómo cambia tu estimación inicial tras aplicar la RPC y cómo afecta esto al umbral de acción.

Tabla 3.- Criterios de calidad metodológica de estudios de evaluación de reglas de predicción clínica

<p>a. Conjunto de variables predictivas válidas, precisas, accesibles y clínicamente importantes (definición clara; reproducibilidad)</p> <p>b. Comparación independiente con un patrón de referencia válido que clasifique el evento, diagnóstico o pronóstico que se quiere predecir (enmascaramiento; definición clara y exhaustiva; importancia clínica)</p> <p>c. Muestra de pacientes adecuada (emplazamiento; criterios de inclusión y exclusión; tamaño muestral suficiente; número de eventos por variable; seguimiento completo)</p> <p>d. Procedimientos estadísticos de diseño y análisis de resultados correctos (técnicas de derivación: análisis multivariante; estimadores de validez: sensibilidad, especificidad, valores predictivos, cocientes de probabilidades, curvas ROC; métodos de calibración)</p> <p>e. Validación prospectiva en nuevas muestras de pacientes (estimadores de validez) y análisis del impacto clínico.</p>

Tabla 4.- Jerarquía de evidencias para las reglas de predicción clínica

<p>Nivel I: Las reglas pueden emplearse en una amplia variedad de contextos con la seguridad de que modificarán el comportamiento médico y mejorarán los resultados de los pacientes</p> <p>En este nivel, las reglas deben haber pasado al menos una validación prospectiva en una población diferente más un análisis del impacto, con la demostración de un cambio en el comportamiento clínico con consecuencias beneficiosas</p>
<p>Nivel II: Las reglas pueden emplearse en una amplia variedad de contextos con una precisión fiable</p> <p>En este nivel, las reglas deben haber mostrado su precisión ya sea en un gran estudio prospectivo con una amplia gama de pacientes y médicos, o bien mediante validación en varios contextos pequeños que difieran los unos de los otros</p>
<p>Nivel III: Las reglas pueden ser empleadas por los médicos con precaución y sólo si los pacientes del estudio son semejantes al del contexto clínico concreto</p> <p>Estas reglas sólo se han validado en una sola muestra prospectiva limitada</p>
<p>Nivel IV: Las reglas necesitan de mayor validación antes de poder aplicarlas clínicamente</p> <p>Estas reglas se han derivado pero no se han validado, o han sido validadas sólo en muestras sesgadas, en grandes bases de datos retrospectivas o mediante técnicas estadísticas</p>