



Uso apropiado de los métodos diagnósticos en la decisión clínica

Juan Gérvas
Médico General.
Equipo CESCA. Madrid.

Palabras clave:

- > Precisión diagnóstica
- > Decisión clínica
- > Impacto en salud

Mercedes Pérez Fernández
Médico General.
Equipo CESCA. Madrid.

INTRODUCCIÓN

Diagnosticar es un arte. Pero lo importante no es el diagnóstico en sí, sino la ayuda que conlleva tener un diagnóstico en el manejo del paciente, en la toma de decisiones, en la estimación del pronóstico y en la respuesta al tratamiento. El médico llega al diagnóstico a través del conocimiento del paciente en su entorno, y para ello utiliza mucha información, formal e informal¹.

Lo básico es la historia clínica y la exploración del paciente, más el conocimiento de la comunidad. Con esta información el médico disminuye probabilidades^{2,3}. Si la probabilidad del diagnóstico es incierta, o conviene calibrar su gravedad, el médico decide solicitar pruebas diagnósticas, con lo que se inicia una cascada con efectos no sólo beneficiosos⁴. De ahí la importante ganancia de eficiencia que conlleva una medicina general o de familia, pues los médicos generales o de familia tienen un acervo increíble de información preprueba, por ofrecer continuidad (seguimiento de los problemas) y longitudinalidad (seguimiento de las personas y de las familias) (1).

El buen ejercicio del médico de familia permite valorar a muy bajo costo probabilidades previas (por el conocimiento de la población, de las familias y de los pacientes, «de ciencia propia»), y aumenta la probabilidad preprueba⁵. En la práctica, según el teorema de Bayes, la probabilidad posprueba de enfermedad depende más de la exactitud con que se haya determinado la probabilidad preprueba de la enfermedad que de las características de la prueba diagnóstica.

Pese a la importancia de los métodos diagnósticos, no existe un proceso de validación normativa de los mismos, similar al impuesto sobre los métodos terapéuticos (ensayos clínicos para evaluar la eficacia, y estudios observacionales para tener información sobre tolerancia y seguridad)⁶. No es de extrañar, pues, una gran variabilidad en su uso⁷. Variabilidad extremadamente

importante, pues los métodos diagnósticos han ampliado su campo y facilitan no sólo el diagnóstico en sí, sino también la toma de decisión respecto al pronóstico, las complicaciones esperables, el tratamiento y la respuesta a éste, la determinación de trastornos subclínicos y el impacto en la salud del paciente^{8,9}.

Bibliografía

1. Tanenbaum S. What physicians know. *N Engl J Med.* 1993;329:1268-71.
2. Latour J. El diagnóstico. *Quaderns de salut pública i administració de serveis de salut*, 21. Valencia: Escola d'Estudis per a la Salut; 2003.
3. Montori VM, Wyer P, Newman TB, Keitz S, Guyatt GH, for the EBMTWG. Tips for learners of evidence-based medicine. The effect of the spectrum of disease on the performance of diagnostic tests. *CMAJ.* 2005;173:385-90.
4. Pérez Fernández M, Gérvas J. El efecto cascada: implicaciones clínicas, epidemiológicas y éticas. *Med Clin (Barc).* 2002;118:65-7.
5. Ortún V, Gérvas J. Fundamentos y eficiencia de la atención médica primaria. *Med Clin (Barc).* 1996;106:97-102.
6. Pijoán JI, Ruiz García V. La compleja relación entre el diagnóstico y la práctica clínica. *Semergen.* 2006;32:1-3.
7. Ripoll Lozano MA, Alda C, Pérez Fernández M, Gérvas J. Estudio de la demanda de pruebas de laboratorio en atención primaria. *Aten Primaria.* 1995;16:73-8.
8. Guyatt GH, Tugwell PX, Feeney DH, Haynes RB, Drummond M. A framework for clinical evaluation of diagnostic technologies. *CMAJ.* 1986;134:587-94.
9. Knottnerus JA, editor. *The evidence base of clinical diagnosis.* Londres: BMJ books; 2002.

> 1

EL BRILLO DE LAS NUEVAS Y ALTAS TECNOLOGÍAS

Pregunta

¿A qué se debe la variabilidad médica?

Respuesta

A la «atención discrecional sensible a las preferencias de los pacientes» y a la «atención sensible a la oferta, a las preferencias de los profesionales».

(1) En este texto utilizamos indistintamente los términos medicina de familia, medicina general, médico de familia y médico general, pues en los artículos comentados los hay de cultura estadounidense (que prefiere médico de familia), y europea (que prefiere médico general).

Resumen del artículo

Peiró S, Bernal-Delgado E. ¿A qué incentivos responde la utilización hospitalaria en el Sistema Nacional de Salud? *Gac Sanit.* 2006;20(Supl 1):110-6.

Contexto. Las variaciones geográficas de la práctica médica son una constante de difícil justificación.

Objetivo. Analizar los factores asociados a la utilización de servicios en el Sistema Nacional de Salud español.

Tipo de estudio. Observacional.

Métodos. Se compararon las tasas de utilización por áreas geográficas para algunos procedimientos hospitalarios en el Sistema Nacional de Salud, y en el Medicare estadounidense (sistema de financiación pública, federal, en Estados Unidos, que contrata servicios privados para atender a todos los ancianos, y a los discapacitados reconocidos como tales). Como indicador de «oferta» se empleó la tasa de camas por 10.000 habitantes.

Resultados. La oferta española no se asocia a mayor utilización de procedimientos que necesitan poca tecnología, tipo herniorrafia inguinal. Tampoco hay la típica asociación estadounidense entre oferta e ingresos por procesos crónicos descompensados (insuficiencia cardíaca, enfermedad pulmonar obstructiva crónica [EPOC], diabetes). Pero cuando existe alta utilización de un procedimiento en un área española, se asocia a gran utilización de muchos de los procedimientos estudiados. Hay, pues, áreas españolas que intervienen más de todo, y áreas que intervienen menos de todo. Las tasas de los procesos que se asocian a innovaciones tecnológicas parecen aproximarse notoriamente a las tasas de la población de Medicare.

Conclusión. Los médicos estadounidenses tendrían por estímulo el maximizar ingresos hospitalarios, mientras los españoles obedecerían más a una hipótesis de «fascinación tecnológica», independiente de los intereses de los pacientes. El brillo tecnológico atraería a los profesionales por cuestiones de prestigio y de estatus.

Comentario

En este trabajo se presentan datos que permiten apuntalar la hipótesis sociológica de que los pacientes «son el combustible del sistema sanitario»¹. Así se deduce de la fascinación tecnológica, que parece llevar a un comportamiento en que se aborrecen las afecciones más habituales, frecuentes y crónicas, en las que se emplean «bajas» tecnologías bien conocidas. Esta hipótesis de la fascinación tecnológica implica que el Sistema Nacional de Salud no se orienta hacia las necesidades de los pacientes, sino que tendría el rumbo marcado por las preferencias individuales o profesionales de los médicos, muy sensibles al brillo de la tecnología. Una tendencia opuesta a lo que sería deseable en el siglo XXI^{2,3}.

Bibliografía

1. Irigoyen J. La crisis del sistema sanitario en España: una interpretación sociológica. Granada: Universidad de Granada; 1996.

2. Tinetti ME, Fried T. The end of the disease era. *Am J Med.* 2004;116:179-85.
3. Gérvas J. Moderación en la actividad preventiva y curativa. Cuatro ejemplos de necesidad de prevención cuaternaria en España. *Gac Sanit.* 2006;20(Supl 1):127-34.

> 2 LA NECESARIA VARIABILIDAD

Pregunta

¿Hay un cierto grado de variabilidad necesaria?

Respuesta

Sí, la necesaria para adaptarse al contexto.

Resumen del artículo

Turabián-Fernández JL, Pérez-Franco B. La variabilidad es un indicador de buena gestión clínica en medicina de familia. *Aten Primaria.* 2006;37:160-3.

Contexto. La media estadística como ideal resulta absurda en medicina general o de familia.

Objetivo. Analizar el fenómeno de la variabilidad en la práctica médica en atención primaria.

Tipo de estudio. Revisión, análisis y opinión.

Métodos. Los autores enfrentan la cantidad a la calidad, y los números y las poblaciones a las situaciones personales. Con ello (y con un docente caso clínico) intentan determinar la existencia de una variabilidad adecuada y razonable.

Resultados. La tarea primera del médico es responder a la experiencia del enfermo, de forma que pueda ofrecer los cursos de acción óptimos para que el propio paciente elija alguno (o rechace todos). Diversos pacientes con situaciones similares pueden y deben recibir intervenciones diagnósticas y terapéuticas diferentes. El intento normalizador de la medicina basada en pruebas tiene un componente negativo cuando pretende disminuir sin más la variabilidad clínica. Sus aplicaciones, a través de guías y protocolos, sólo pueden emplearse después de tener en cuenta el contexto. La enfermedad es un sistema complejo y adaptativo, no el fallo de un sistema simple.

Conclusión. La naturaleza nos brinda el ejemplo de la biodiversidad como estrategia de evolución, y la variabilidad es sinónimo de vida. Por ello es esperable y deseable un cierto grado de variabilidad en atención primaria, donde hay vida. Si la variabilidad es menor de la requerida, el proceso de atención será inadecuado. Esta variabilidad está bien lejos del «brillo tecnológico», de la variabilidad no justificada que se debe al prestigio de la nueva y alta tecnología.

Comentario

Existe una enorme y preocupante variabilidad en las recomendaciones y prácticas fundadas en la medicina basada en pruebas¹. Así pues, incluso si se utilizasen las guías y protocolos fundados en pruebas, lo lógico sería esperar una alta variabilidad, por la multitud de dichos instrumentos. ¿Cómo no esperar variabilidad, si además estamos ante un sistema complejo, como la medicina general²? La práctica médica está llena de interrogantes sin respuestas definitivas³.

Los datos de que disponemos sobre la salud y la enfermedad son incompletos, insuficientes, y muchas veces erróneos. La aplicación no contextualizada de los resultados de los ensayos clínicos, procedentes de «poblaciones de pacientes», no puede hacerse sin el riesgo de errar (y de hacer daño) en los individuos. Este problema ha sido denominado «tragedia clínico-estadística»⁴. Un cierto grado de variabilidad es saludable e indicador de buen trabajo clínico en atención primaria.

Bibliografía

1. Steinberg EP, Luce BR. Evidence-based? Caveat emptor? Health Affairs. 2005;24:80-92.
2. Love T, Burton C. General practice as a complex system: a novel analysis of consultation data. Fam Pract. 2005;22:347-52.
3. Gérvas J. Experiencia, ciencia y «medicina basada en pruebas» en atención primaria. Semergen. 2002;28:302-4.
4. Feinstein AR. The problem of cogent subgroups: a clinico-statistical tragedy. J Clin Epidemiol. 1998;51:297-9.

> 3

LA VARIABILIDAD Y LA MALA CALIDAD DE LOS ESTUDIOS SOBRE MÉTODOS DIAGNÓSTICOS

Pregunta

¿Son de fiar los estudios sobre precisión y validez de los métodos diagnósticos?

Respuesta

Los estudios sobre métodos diagnósticos son de baja calidad.

Resumen del artículo

Rutjes AWS, Reitsma JB, Di Nisio M, Smidt N, Rijn JCV, Bossuyt PMM. Evidence of bias and variation in diagnostic accuracy studies. CMAJ. 2006;174:469-76.

Contexto. La precisión se refiere al grado de concordancia entre la información ofrecida por la prueba diagnóstica evaluada y la ofrecida por la prueba de referencia (*gold standard*).

Objetivo. Valorar la calidad de los estudios sobre precisión de pruebas diagnósticas a partir de los metaanálisis que los revisan.

Tipo de estudio. Revisión sistemática de la literatura científica.

Métodos. Revisión de todos los metaanálisis publicados sobre precisión de pruebas diagnósticas (1999-2002) en alemán, francés, holandés e inglés. Seleccionaron los metaanálisis que al menos incluyeran 10 estudios originales. Como medida de precisión de la prueba diagnóstica emplearon la *odds ratio* diagnóstica (un índice que sintetiza la sensibilidad y la especificidad).

Resultados. Identificaron un total de 191 revisiones sistemáticas, de las que pudieron utilizar 31 metaanálisis sobre 487 estudios de precisión de pruebas diagnósticas. Entre los 487 trabajos publicados sólo uno cumplió los criterios de calidad del cuestionario elaborado. La mayor sobreestimación de la *odds ratio* diagnóstica se dio en los trabajos en los que se midió la precisión de la prueba comparando los resultados en sanos frente a enfermos, pues se multiplicó casi por cinco. La mayor infraestimación se dio en aquellos en los que se midió la precisión de la prueba en pacientes derivados, pues fue de la mitad. En general, la calidad de los estudios sobre la precisión de las pruebas diagnósticas es bajísima, lo que dificulta la evaluación y valoración de las publicaciones sobre ellos.

Conclusión. Son frecuentes los sesgos y la variabilidad en la evaluación de la precisión de las pruebas diagnósticas. Las publicaciones sobre pruebas diagnósticas son de muy baja calidad, lo que dificulta su análisis.

Comentario

La búsqueda heroica del resultado anormal es una actividad carente de ciencia y conciencia¹. Las sucesivas evaluaciones de las pruebas diagnósticas, a partir de la más clásica de Lijmer et al², han demostrado la mala calidad de los estudios acerca de la precisión de las mismas. Así pues, el arsenal de pruebas diagnósticas aumenta sin cesar, como consecuencia de los avances de la técnica, pero no aumenta ni la calidad ni el rigor de la evaluación de su precisión o exactitud diagnóstica³. Tampoco se estudian con rigor otras cuestiones de gran interés clínico, como el impacto de la prueba en la salud y el manejo futuro del caso. ¿Es esperable un uso adecuado de pruebas cuya utilidad y calidad desconocemos? ¿Es el brillo tecnológico, de nuevo, el que guiará la actuación del clínico, y la expectativa del paciente?

Bibliografía

1. Woolf S, Kamerov DB. Testing for uncommon conditions. The heroic search for positive test results. Arch Intern Med. 1990;150:2451-8.
2. Lijmer JG, Mol BW, Helsterekamp S, Bossel GJ, Prins MH, Meulen GHV, et al. Empirical evidence of design-related bias in studies of diagnostic tests. JAMA. 1999;282:1061-6.
3. Pijoán JI, Ruiz-García V. La compleja relación entre el diagnóstico y la práctica clínica. Semergen. 2006;32:1-3.

> 4

LOS SESGOS NUNCA VIENEN SOLOS

Pregunta

¿Cuáles son los errores más frecuentes en la estimación de la precisión y validez de las pruebas diagnósticas?

Respuesta

La selección de los pacientes, la independencia entre la prueba y el patrón de referencia, y la falta de enmascaramiento.

Resumen del artículo

Abraira V. Sesgos en los estudios sobre pruebas diagnósticas. *Semergen*. 2006;32:24-6.

Contexto. Cuando se trata de valorar la precisión y validez de las pruebas diagnósticas no es infrecuente la ausencia de información relevante.

Objetivo. Analizar los sesgos más frecuentes en el diseño de los estudios sobre precisión de pruebas diagnósticas.

Tipo de estudio. Revisión, análisis y opinión.

Métodos. El autor repasa los aspectos del diseño de los estudios sobre la precisión de las pruebas diagnósticas, analiza los sesgos más frecuentes en los mismos y cuantifica empíricamente el impacto de dichos sesgos en la interpretación de la validez de las pruebas diagnósticas.

Resultados. En el diseño ideal de los estudios de evaluación de la validez de pruebas diagnósticas, éstas se aplican al tipo habitual de paciente en el que se utilizará en el futuro, y se comparan los resultados de la prueba índice con los de la prueba de referencia (realizada simultáneamente). Esta comparación debería ser enmascarada. El sesgo más frecuente se da en el diseño de casos control, al comparar los resultados en un grupo de pacientes que tienen la enfermedad con los resultados en un grupo de sanos. También es frecuente que las pruebas índice y de referencia no sean independientes, lo que da lugar a un sesgo de verificación parcial. Por último, se da un sesgo importante cuando no se hace a ciegas la comparación entre los resultados de la prueba índice y la de referencia.

Conclusión. Aunque el diagnóstico tiene un papel central en la actividad médica, los médicos tienen escaso entrenamiento formal sobre la utilización de las pruebas diagnósticas, y sobre la mejor alternativa para formarse una opinión sólida respecto al valor de las nuevas pruebas que se introducen de continuo en la práctica clínica.

Comentario

La mejor aproximación clínica se logra a través de un diagnóstico diferencial más completo y ponderado, una selección más eficiente de las pruebas complementarias, y una integración rápida del conjunto de la información en la resolución del problema diagnóstico del paciente concreto. Necesitamos conocer más este proceso diagnóstico, tan característico de los médicos pero tan mal entendido^{1,2}. Además, el diagnóstico es un primer paso clave, que lleva inevitablemente, en una cascada de gran importancia clínica, hasta la terapéutica^{3,4}.

En el diagnóstico, más tecnología no siempre es mejor, especialmente cuando desconocemos su utilidad.

Bibliografía

1. Guyatt GH, Tugwell PX, Feeny DH, Haynes RB, Drummond M. A framework for clinical evaluation of diagnostic technologies. *CMAJ*. 1986;134:587-94.
2. Latour J. El diagnóstico. *Quaderns de salut pública i administració de serveis de salut*, 21. Valencia: Escola d'Estudis per a la Salut; 2003.
3. Pérez Fernández M, Gervas J. El efecto cascada: implicaciones clínicas, epidemiológicas y éticas. *Med Clin (Barc)*. 2002;118:65-7.
4. Anónimo. Como una piedra rodando (*like a rolling stone*). *Informatiu AATRM*. 2006;38:1-3. Disponible en: <http://www.gencat.net/salut/depsan/units/aatrm/pdf/but38ca.pdf>

> 5

UNA PROPUESTA PARA DISMINUIR LA VARIABILIDAD DE LOS ESTUDIOS SOBRE PRUEBAS DIAGNÓSTICAS

Pregunta

¿Hay forma de disminuir la variabilidad de los estudios sobre precisión de pruebas diagnósticas?

Respuesta

Sí, mediante el establecimiento de directrices homogéneas.

Resumen del artículo

Altman DG, Bossuyt PMM. Estudios de precisión diagnóstica (STARD) y pronóstica (REMARK). *Med Clin (Barc)*. 2005 (Supl 1):49-55.

Contexto. Ha habido intentos varios de establecer directrices homogéneas, del estilo de las que se han impuesto para los ensayos clínicos.

Objetivo. Describir las iniciativas desarrolladas para el establecimiento de directrices de comunicación sobre estudios de precisión diagnóstica y pronóstica de pruebas diagnósticas.

Tipo de estudio. Revisión, análisis y opinión.

Métodos. El trabajo describe el proceso de desarrollo de los puntos incluidos en las directrices para publicar estudios sobre precisión diagnóstica (STARD, Standards for Reporting of Diagnostic Accuracy) y pronóstica (REMARK, Reporting Recommendations for Tumour Marker Prognostic Studies).

Resultados. La lista de elementos STARD para la publicación de estudios relativos a la precisión diagnóstica tiene 25 puntos. Va desde el título (se recomienda que aparezcan las palabras «sensibilidad y especificidad») hasta la discusión en el artículo (se recomienda que se comente la utilidad de los resultados del estudio en la práctica clínica).

Las recomendaciones para la publicación de estudios sobre el valor pronóstico de los marcadores tumorales (REMARK) contienen 20 elementos desde la introducción (se recomienda, p. ej., que se especifique toda hipótesis inicial) a la discusión (se recomienda la interpretación de los resultados según la hipótesis inicial, y según otras hipótesis alternativas).

Conclusión. Con las listas sólo se busca una mejora de las publicaciones, pero es de esperar que ello conlleve un mejor diseño de los estudios, y una más fácil interpretación por los médicos clínicos de sus resultados.

Comentario

Es lamentable la calidad de las publicaciones sobre la precisión de las pruebas diagnósticas. Por ejemplo, a lo largo de los años se han publicado millares de estudios sobre marcadores tumorales en oncología, aunque el resultado final ha confirmado la utilidad de sólo de unos pocos¹. En otro campo distinto, la utilidad clínica de los «marcadores genéticos» (la capacidad predictiva de genes diagnósticos), de 600 genes atribuidos a distintos problemas de salud, sólo 6 han confirmado tal atribución al cabo del tiempo².

Parece, pues, que en el diagnóstico el deseo va por delante de las posibilidades de la tecnología; o, mejor, que la tecnología encandila con su brillo y lleva a sesgos repetidos que sobrevaloran el valor de las pruebas diagnósticas³.

Bibliografía

1. Schilsky RL, Taube SE. Tumor markers as clinical cancer tests: are we there yet? *Semin Oncol.* 2002;29:211-1.
2. Hirschorn HN, Lohmueller K, Byrne E, Hirschorn K. A comprehensive review of genetic association studies. *Gen Med.* 2002;4:45-61.

(2) La sensibilidad y la especificidad son características que tienen en cuenta el carácter de enfermo o sano del paciente (permiten clasificar a la población de pacientes), y no indican la capacidad de la prueba diagnóstica para confirmar la enfermedad cuando el resultado es positivo, o negativo. Es decir, la sensibilidad y la especificidad nos hablan de la capacidad de la prueba para detectar enfermos, no de la probabilidad de enfermedad posprueba. Los cocientes de probabilidad, o de verosimilitud, ayudan a valorar la probabilidad posprueba, a partir de la probabilidad preprueba, y dependen del punto de corte a partir del cual definimos la prueba como positiva o negativa.

3. Whiting P, Rutjes AW, Reitsma JB, Glas AS, Bossuyt PM, Kleijnen J. Sources of variation and bias in studies of diagnostic accuracy: a systematic review. *Ann Intern Med.* 2004;140:189-202.

> 6 LOS MÉDICOS APLICAN APROPIADAMENTE EL TEOREMA DE BAYES

Pregunta

¿Es mejor expresar la precisión de las pruebas diagnósticas como sensibilidad y especificidad, o como los cocientes de probabilidad? (2).

Respuesta

Los médicos interpretan sin dificultad el impacto de la información del resultado de una prueba diagnóstica.

Resumen del artículo

Puhnan MA, Stenrer J, Bachmann LM, Riet G. A randomized trial of ways to describe test accuracy: the effect on physicians' post-test probability estimates. *Ann Intern Med.* 2005;143:184-9.

Contexto. El resultado de la prueba modifica la probabilidad preprueba de que el paciente tenga (o no tenga) la enfermedad que se quiere confirmar (o descartar).

Objetivo. Comprobar si los médicos se «acomodan» al teorema de Bayes, con independencia de los términos con que se les presenta la información.

Tipo de estudio. Ensayo clínico aleatorizado.

Métodos. Se prepararon 6 casos clínicos en que se consideraba la probabilidad de un diagnóstico de acuerdo con la información clínica (preprueba), y tras recibir el resultado positivo anormal de una prueba diagnóstica acorde con el caso (posprueba). Los casos fueron de probable infarto de miocardio, embolia pulmonar, EPOC, arteritis temporal, insuficiencia cardíaca y gripe. Los datos sobre la precisión de la prueba se dieron: a) como sensibilidad y especificidad (en porcentajes), b) como cocientes de probabilidad numéricos en lenguaje no técnico, y c) como cocientes de probabilidad, de forma gráfica.

Resultados. Al calcular las probabilidades posprueba, la tendencia de los médicos fue a cambiar menos las probabilidades de enfermedad de lo que permite el teorema de Bayes, con las tres formas de presentación de la precisión de la prueba diagnóstica considerada. La mitad de los errores se dieron en los casos en que la prueba diagnóstica tiene baja precisión (embolia pulmonar y arteritis de la temporal).

Conclusión. Los médicos aplican el teorema de Bayes con propiedad, para cambiar probabilidades preprueba tras conocer los resultados de las pruebas diagnósticas (probabilidades postprueba). En general son conservadores al respecto, y los errores se producen cuando las pruebas diagnósticas tienen poca precisión.

Comentario

En el médico no influye el hablar de sensibilidad y especificidad o de cocientes de probabilidad. Lo que parece influir es el caso clínico, de forma que la probabilidad preprueba cambia poco tras recibir la prueba (cambia menos de lo que permitiría el teorema de Bayes), en una actitud conservadora que discrepa de la sobreestimación que puede producir el resultado de una sola prueba, por muy informativa que sea.

El contexto es muy importante, tanto para el médico como en la teoría de la decisión^{1,2}. En muchos aspectos, las simplificaciones teóricas del diagnóstico clínico no consiguen explicar de forma satisfactoria la complejidad de las situaciones que resuelven acertadamente los médicos³. Un cierto grado de variabilidad es, pues, una característica de la buena atención clínica.

Bibliografía

1. Miettinen OS. The modern scientific physician. Scientific diagnosis. CMAJ. 2001;165:781-2.
2. Moons KG, Es GAW, Deckers JW, Habbena JD, Grobbee DE. Limitations of sensitivity, specificity, likelihood ratio, and Bayes' theorem in assessing diagnostic probabilities: a clinical example. Epidemiology. 1997;8:12-7.
3. Reid CM, Lane DA, Feinstein AR. Academic calculations versus clinical judgments: practising physicians' use of quantitative measures of test accuracy. Am J Med. 1998;104:374-80.

> 7

LA CERTEZA PARA EXCLUIR EL MAL PRONÓSTICO DE LA EMBOLIA PULMONAR

Pregunta

¿Es posible seleccionar a un grupo de pacientes altamente sospechosos de presentar embolia pulmonar con baja probabilidad de morir?

Respuesta

Sí, con diez datos clínicos objetivos a la cabecera de la cama.

Resumen del artículo

Aujesky D, Scott D, Stone RA, Auble TE, Perrier A, Cornuz J, et al. A prediction rule to identify low-risk patients with pulmonary embolism. Arch Intern Med. 2006;166:169-75.

Contexto. Es importante encontrar métodos diagnósticos objetivos que permitan clasificar a los pacientes con sospecha clínica de embolia pulmonar según su probabilidad de complicaciones mortales.

Objetivo. Elaborar una tabla de predicción (3) basada en hechos clínicos simples y objetivos.

Tipo de estudio. Deductivo, retrospectivo (desarrollo y validación interna) y prospectivo (validación externa).

Métodos. Se eligió una muestra al azar entre 15.531 pacientes hospitalarios con embolia pulmonar demostrada por el paso del tiempo. Con la tabla de predicción elaborada, se validó internamente. Posteriormente se validó la misma tabla en un estudio prospectivo sobre casos hospitalarios de embolia pulmonar en el que la certeza del diagnóstico se logró con el uso sistemático de tomografía computarizada helicoidal.

Resultados. Elementos del modelo final: *a)* edad (mayor o igual a 70 años), *b)* historia de cáncer, *c)* de insuficiencia cardíaca, *d)* de EPOC, *e)* de insuficiencia renal crónica, *f)* de enfermedad cerebrovascular, *g)* taquicardia (de 110 lat/min o mayor), *h)* hipotensión sistólica (de 100 mmHg o menos), *i)* presencia de alteración del estado mental, *j)* saturación arterial de oxígeno menor del 90%. Los pacientes que no cumplían estos criterios tuvieron una mortalidad en los 30 días siguientes al diagnóstico del 0,6% en la muestra inicial, del 1,5% en la muestra de validación interna y del 0% en la muestra de validación externa.

Conclusión. Los 10 datos seleccionados permiten diferenciar a un importante grupo de pacientes (en torno al 35%) que aun teniendo embolia pulmonar evolucionarán sin graves problemas, y pueden controlarse en el domicilio.

Comentario

Si contamos con una tabla pronóstica a partir de elementos simples y objetivos, podemos obviar el llegar a la certeza del diagnóstico, y pasar directamente al objetivo que cuenta, el manejo apropiado del paciente¹. Con ello podemos cambiar el lugar del seguimiento del paciente, del hospital a su domicilio (y el médico general puede sustituir al especialista hospitalario, consecuentemente).

Así pues, esta tabla pronóstica del paciente con sospecha o certeza de embolia pulmonar es un avance tecnológico que

(3) Algunas tablas de predicción se conocen también como «tablas de riesgo» o «tablas de estratificación por riesgo». La palabra *riesgo* resulta equívoca, y por ello preferimos el término predicción, más apropiado. Las tablas de predicción propiamente dichas no son «tablas de decisión», pues sólo sirven para determinar probabilidades preprueba diagnósticas o pronósticas, y una cosa es la probabilidad de un evento y otra la decisión clínica. Además, las tablas de predicción no suelen estar estudiadas ni validadas para su uso en la decisión.

puede ayudar a llevar los cuidados más cerca del hogar del paciente, lo que aumenta la seguridad y disminuye la variabilidad y los errores^{2,3}.

Bibliografía

1. Moores LK. Diagnosis and management of pulmonary embolism. Are we moving toward an outcome standard?. *Arch Intern Med.* 2006; 166:147-8.
2. Gervas J, Pérez Fernández M. ¿Cuál es el límite en la prestación de servicios cercanos al paciente? El límite lo marca la tensión entre el síndrome del barquero y el síndrome del gato. *Med Clin (Barc).* 2005; 124:778-81.
3. Gervas J. Moderación en la actividad preventiva y curativa. Cuatro ejemplos de necesidad de prevención cuaternaria en España. *Gac Sanit.* 2006;20(Supl 1):127-34.

> 8

LA ESTRATIFICACIÓN DE LA PROBABILIDAD DE EMBOLIA PULMONAR

Pregunta

¿Es posible estratificar a los pacientes con sospecha de tener embolia pulmonar con criterios clínicos sencillos y objetivos?

Respuesta

Sí, con el índice de Ginebra modificado.

Resumen del artículo

Le Gal G, Righini M, Roy PM, Sánchez O, Aujesky D, Bounameaux H, et al. Prediction of pulmonary embolism in the emergency department: the revised Geneva Score. *Ann Intern Med.* 2006;114:165-71.

Contexto. Es difícil calcular la probabilidad preprueba de forma sencilla y objetiva en los pacientes con sospecha de embolia pulmonar.

Objetivo. Elaborar un índice simple para clasificar la probabilidad preprueba de tener embolia pulmonar.

Tipo de estudio. Deductivo, retrospectivo (desarrollo y validación interna) y prospectivo (validación externa).

Métodos. La selección de los datos para elaborar el índice se hizo a partir de 965 pacientes, en urgencias hospitalarias. El índice de Ginebra modificado se validó internamente en la misma muestra de que se derivó, y externamente con

los pacientes de otro estudio con 756 pacientes, en el que además del proceso diagnóstico se controló a los pacientes durante 3 meses para mejorar la certeza del diagnóstico.

Resultados. El índice va de 0 a 22 puntos: *a)* edad de 65 años o mayor (1 punto), *b)* embolia previa, de extremidades o pulmonar (3 puntos), *c)* cirugía o fractura en el mes previo (2 puntos), *d)* cáncer activo (2 puntos), *e)* dolor espontáneo unilateral en extremidad inferior (3 puntos), *f)* hemoptisis (2 puntos), *g)* taquicardia (3 puntos si está entre 75 y 94 pulsaciones, y 5 puntos si hay más de 94 pulsaciones por minuto), *h)* dolor a la palpación venosa de una extremidad inferior y edema en la misma (4 puntos). Hay baja probabilidad de embolia pulmonar con 0 a 3 puntos, probabilidad intermedia con 4 a 10 puntos, y probabilidad alta con 11 o más puntos. En la muestra de validación, hubo embolia pulmonar en el 8, el 28 y el 74% de los pacientes con baja, intermedia y alta probabilidad preprueba, respectivamente.

Conclusión. Para clasificar la probabilidad preprueba de tener embolia pulmonar los datos necesarios proceden de la historia y de la exploración física.

Comentario

Se han desarrollado dos índices para facilitar la selección de pacientes respecto a su probabilidad preprueba de tener una embolia pulmonar, el canadiense de Wells, y el suizo-francés de Ginebra^{1,2}.

El objetivo del índice de Ginebra modificado no es tanto ser una ayuda diagnóstica sino ser una ayuda a la decisión. Con una baja probabilidad preprueba de embolia pulmonar, vale la pena utilizar sólo técnicas no invasivas (p. ej., determinación del dímero D) para descartar el diagnóstico. Con una alta probabilidad preprueba se justifica el empleo de técnicas agresivas e invasivas y el tratamiento anticoagulante.

Podemos clasificar a los pacientes con sospecha de embolia pulmonar, utilizando sólo criterios clínicos a la cabecera de la cama, en: *a)* los que no tendrán mala evolución (con independencia de la presencia o ausencia de embolia), y *b)* los que podrán tener mala evolución, agrupados por grado de probabilidad preprueba diagnóstica (baja, media, alta). Todo ello es posible desde la atención primaria, como paso inicial o final de tratamiento de estos pacientes.

Bibliografía

1. Wells PS, Anderson DR, Rodger M, Ginsberg JS, Kearon C, Gent M, et al. Derivation of a simple clinical model to categorize patients probability of pulmonary embolism: increasing the models utility with the SimpiRED D-dimer. *Thromb Haemost.* 2000;83:416-20.
2. Wicki J, Perneger TV, Junod AF, Bounameaux H, Perrier A. Assessing clinical probability of pulmonary embolism in the emergency ward: a simple score. *Arch Intern Med.* 2001;161:92-7.

> 9 ESTUDIOS DE IMPACTO

Pregunta

¿Puede transformarse una tabla de predicción en una tabla de decisión?

Respuesta

Sí, a través de los estudios de impacto.

Resumen del artículo

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.

Contexto. Muchas veces, las tablas de predicción se aplican erróneamente como tablas de decisión.

Objetivo. Revisar las formas de desarrollo de las tablas de predicción y de decisión.

Tipo de estudio. Revisión, análisis y opinión.

Métodos. Los autores definen con precisión que una tabla de decisión se refiere al proceso de atención respecto a un aspecto concreto de un problema de salud.

Resultados. Los cinco pasos para llegar a una tabla de decisión son: *a)* derivación de la tabla de predicción, *b)* validación inicial de la tabla de predicción, *c)* validación general de la tabla de predicción, *d)* estudio de impacto inicial y *e)* estudio de impacto general. Por impacto se entiende tanto repercusión en el uso adecuado de los recursos (resultado intermedio) como en el cambio en el estado de salud del paciente (resultado final). A los médicos les interesa especialmente el valor predictivo negativo (que los pacientes excluidos no tengan el evento considerado), y la seguridad (que todos los que lo precisan reciban los cuidados necesarios), mientras a los gerentes y políticos les interesa el valor predictivo positivo (que los pacientes incluidos tengan el evento considerado) y la eficiencia (definida aquí como que los pacientes que no necesitan cuidados nos los reciban).

Conclusión. Las tablas de predicción pueden transformarse en tablas de decisión a través de estudios de impacto.

Comentario

El manejo adecuado de la incertidumbre es la cruz de la Medicina de Familia. No es extraño, así, que las tablas de predicción se transformen en tablas de riesgo, e insensiblemente acaben siendo tablas de decisión (sobre el manejo diagnóstico y/o terapéutico del paciente). De esta forma disminuye la

incertidumbre y el médico general se siente seguro en la consulta, con la falsa tranquilidad de estar tomando decisiones «basadas en la evidencia», lo que es un error, como ya advierte la «tragedia clínico-estadística»¹ y como se ha demostrado en nuestra práctica².

Las tablas de decisión deberían desarrollarse con el ánimo de incrementar la tasa de verdaderos positivos sin disminuir la tasa de verdaderos negativos (4) (así se complementan los intereses de médicos y de gerentes).

En la práctica, el médico siempre tenderá a aumentar la seguridad a costa de la eficiencia (como se definen en este texto). Por ello, el impacto real de las tablas de decisión será mayor que el impacto potencial en lo que respecta a la seguridad, y menor en lo referente a la eficiencia. Es decir, recibirán el seguimiento y tratamiento adecuado más pacientes que lo necesitan a costa de aplicar seguimiento y tratamiento innecesario a más pacientes que no lo necesitan.

Bibliografía

1. Feinstein AR. The problem of cogent subgroups: a clinicalstatistical tragedy. *J Clin Epidemiol.* 1998;51:297-9.
2. Jimeno J, Molist N, Franch J, Serrano V, Serrano L, Gracia R. Variabilidad en la estimación del riesgo coronario en la diabetes mellitus tipo 2. *Aten Primaria.* 2005;35:30-6.

> 10 MÁS CALIDAD, MENOS CANTIDAD, MÁS CERCA DEL DOMICILIO DEL PACIENTE (Y CON LA VARIABILIDAD NECESARIA)

Pregunta

¿Son los trabajos sobre embolia pulmonar ejemplos acerca de pruebas diagnósticas?

Respuesta

Sí, son un buen ejemplo.

Resumen del artículo

Sox HC. Better care for patients with suspected pulmonary embolism. *Ann Intern Med.* 2006;144:210-2.

(4) En un ejemplo clásico, la sociedad es más tolerante con los falsos positivos de apendicitis (pacientes intervenidos en falso por apendicitis) que con los falsos negativos de apendicitis (pacientes con apendicitis y retrasos diagnósticos, que pueden desarrollar complicaciones graves, como peritonitis, o morir).

Contexto. Los avances prácticos en el campo del proceso diagnóstico no se han acompañado del lógico desarrollo de validación normativa.

Objetivo. Valorar en clave general los estudios sobre el diagnóstico de embolia pulmonar.

Tipo de estudio. Revisión, análisis y opinión.

Métodos. El autor, director del *Annals of Internal Medicine*, comenta tres artículos publicados en un mismo número de la revista acerca del proceso diagnóstico (5).

Resultados. Los médicos están más dispuestos a seguir las «normas» cuando se trata de confirmar el diagnóstico de embolia pulmonar (92%) que cuando se trata de descartar el mismo diagnóstico (57%). Parece que la mejor explicación es que los médicos hacen una estimación de la probabilidad preprueba muy baja, y/o que sobreestiman en exceso el resultado negativo de las pruebas diagnósticas cuando tratan de descartar el diagnóstico. Los índices de Wells y de Ginebra pueden ayudar a corregir esa infraestimación de la probabilidad preprueba, y el uso del teorema de Bayes puede ayudar a reducir la sobreestimación del resultado negativo de las pruebas. Todo ello está impregnado de sentido común. El índice de Ginebra modificado ayuda como el previo a calcular la probabilidad preprueba del diagnóstico de embolia pulmonar, pero además lo hace con datos de la historia y de la exploración física. Ello supone un avance inconmensurable, si se valida externamente en otros ámbitos, como atención primaria. Puede llevar el manejo del paciente al domicilio del mismo, sobre todo si se combina con la tabla de riesgo de muerte por embolia pulmonar.

Los estudios de impacto (de tratamiento de pacientes) permiten transformar las tablas de predicción en tablas de decisión. En estos estudios se pueden analizar las conductas diagnósticas de los médicos, y su mayor énfasis en la seguridad (tratar a todos los pacientes que hay que tratar) que en la eficiencia (no tratar a todos los pacientes que no hay que tratar).

Conclusión. Los trabajos sobre el diagnóstico de la embolia pulmonar han abierto nuevos caminos.

Comentario

Al médico y al paciente les rebela el que no se pueda hablar categóricamente en medicina. Por ejemplo, en lo que respecta al diagnóstico, casi siempre se trata de posibilidades. Además, toda prueba diagnóstica abre una cascada de final impredecible, y tiene cierta agresividad (que aumenta conforme se quiere disminuir la incertidumbre). Y toda nueva prueba añade mayor coste, para el paciente y para la sociedad, en tiempo, colas y gasto. Por otra parte, no conviene llevar al límite la capacidad diagnóstica, por el riesgo del sobretratamiento. Por ejemplo, los nuevos aparatos de tomografía computarizada pueden llegar a demostrar microcoágulos en la circulación pulmonar de escasa o nula importancia clínica, lo que sólo añade el riesgo de un tratamiento innecesario¹.

Es prudente aprovechar la estela de lo publicado sobre embolia pulmonar, que ayuda tanto al análisis de dicho proceso al tiempo que abre campos muy cercanos a los intereses de la medicina general, como el lugar apropiado para la prestación de servicios^{2,3}. Todo ello incide en la máxima con la que debería trabajar el buen clínico: «máxima calidad, mínima cantidad, tan cerca del paciente como sea posible» (y con la variabilidad necesaria)⁴.

Bibliografía

1. Le Gal G, Righini M, Roy PM, Sánchez O, Aujesky D, Bounameaux H, et al. Prediction of pulmonary embolism in the emergency department: the revised Geneva Score. *Ann Intern Med.* 2006;144:165-71.
2. Gervas J, Pérez Fernández M. ¿Cuál es el límite en la prestación de servicios cercanos al paciente? El límite lo marca la tensión entre el síndrome del barquero y el síndrome del gato. *Med Clin (Barc).* 2005;124:778-81.
3. Moores LK. Diagnosis and management of pulmonary embolism. Are we moving toward an outcome standard? *Arch Intern Med.* 2006;166:147-8.
4. Gervas J. Moderación en la actividad preventiva y curativa. Cuatro ejemplos de necesidad de prevención cuaternaria en España. *Gac Sanit.* 2006;20(Supl 1):127-34.

(5) Dos de estos artículos son los comentados en esta revisión con los números 8 y 9.