

Notas metodológicas

De la evidencia científica a las decisiones clínicas

Estrategia diagnóstica: ¿pescar con redes de arrastre o palangre?

5/03/2006

*Martín Caicoya Gómez-Morán**Jefe de Servicio de Prevención de Riesgos Laborales del Principado de Asturias. Hospital Monte Naranco.*

En 1981 España, y en particular la salud pública, tenía aspecto de crisálida. Salía de un largo letargo y aun no sabía cómo enfrentarse a ese mundo frenético que estaba ocurriendo fuera. Entonces estalló la bomba. Al principio se llamó neumonía atípica, en consonancia con la emergencia de este tipo de neumonías, producidas por una serie de gérmenes que respondía bien a algunas familias de antibióticos. El ministro de sanidad buscó consejo en el desconcierto de su gabinete y de las reuniones sacó conclusiones disparatadas: la epidemia la producía un bichito delicado que si se caía se rompía. Esa manifestación del ministro demostraba cuán desorganizado estaba su departamento.

El número de casos crecía logarítmicamente, llegando diagnosticarse en un solo día 1800 y la respuesta a antibióticos era nula. Estaba claro que se trataba de otra enfermedad. No se sabía ni la causa ni cómo manejarlas. Había que realizar una investigación. Acudieron a los CDCs.

Los CDC (Center for Disease Control and Prevention) se han especializado en la vigilancia y control epidemiológico. Son, en cierta manera, la referencia mundial en este campo. La idea de pedir ayuda a este organismo era buena. Enviaron un epidemiólogo que hablaba español. Recuerdo a Rigau-Perez en televisión explicando cómo iba a realizar la investigación: ante la falta de hipótesis sobre la causa realizaría un estudio comparando una serie de casos con sujetos sin enfermedad respecto a una variedad larga de posibles agentes. Es lo que se llama en epidemiología, y no acabo de saber por qué "una expedición a pescar". En sus estudios, hablo de memoria pues nunca fueron publicados, se encontraban asociaciones inverosímiles, simplemente por casualidad.

Piense que si hacen muchas comparaciones, en alguna se encontrará una asociación simplemente porque una de cada veinte veces que decimos que una característica es más (o menos) frecuente en un grupo que en otro nos equivocamos: aceptamos errar el 5%. Es lo que se llama error tipo I. El error tipo II es la otra cara de la moneda: no habiendo encontrado que la característica sea más (o menos) frecuente en una población que en otra, decimos que esas dos poblaciones son la misma, cuando en realidad no lo son y lo que nos faltó fue capacidad discriminatoria, el llamado poder. Esto suele pasar cuando la muestra es pequeña para encontrar esa diferencia en la distribución de la variable, o cuando se mide con mucho error.

Pues bien, Rigau-Pérez encontró asociaciones poco creíbles. Lanzó una red fina para coger todo lo que había. Hay veces que no queda otro remedio que actuar así. En situaciones de desconcierto, las "expediciones a pescar" pueden servir para formular una hipótesis. Con ella se puede proponer una pesca con caña, eligiendo el cebo adecuado.

Mientras tanto, por las mismas fechas, un pediatra, el Dr. Tabuenca, del Hospital del Niño Jesús, observó que sólo había un caso de un niño lactante con el síndrome. Sin ser precisamente epidemiólogo, pensó en la triada en la que se basa la epidemiología descriptiva: persona, tiempo y lugar. Casi sólo con esas armas ya había descubierto Goldberger a principios del mismo siglo la causa de la epidemia de pelagra: tampoco la sufrían los lactantes. A Goldberger se le ocurrió que algo faltaba en la dieta y que eso era lo que producía la enfermedad. Era la primera vez que se pensaba que una ausencia podía ser una causa. Tabuenca también pensó en la dieta cuando descubrió que el único lactante que tenía la enfermedad estaba siendo alimentado con dieta casi normal. Además observó que los casos ocurrían en familias y en ciertas zonas de Madrid (lugar). Con estas ideas estudió la dieta de 62 niños con la enfermedad y de 62 sin la enfermedad. Encontró que el 100% de los que estaban enfermos habían consumido un aceite comprado en garrafas de 5 litros a vendedores ambulantes. El estudio se hizo entre el 6 y el 8 de junio de 1981. El 10, pasado un mes y 10 días desde el inicio de la epidemia, se anunciaba en televisión que la posible causa era el aceite de colza que se vendía en garrafas. Con algunas reservas, el pequeño estudio de Tabuenca había servido para descubrir el vehículo causal y frenar la progresión de la epidemia. A partir de ese día la curva de incidencia inicia un rápido descenso que es fácil atribuir al anuncio.

Esta investigación no se diferencia mucho del acto médico. En él también se maneja información, para realizar un diagnóstico, establecer un tratamiento que teóricamente modifique (idealmente, que mejore) el curso natural de la enfermedad y aventurar un pronóstico. La pregunta que uno se hace, delante de un paciente que plantea una duda diagnóstica, es qué pruebas pedir y en qué orden. En otras palabras, qué estrategia seguir, la de la pesca de arrastre o la de caña con cebo específico. Cada una de ellas tiene sus ventajas e inconvenientes y es más o menos pertinente según las circunstancias.

Si uno tiene tiempo, porque la enfermedad no progresa rápidamente y se puede dilatar en el tiempo la

realización del diagnóstico, sin costes importantes, es posible que piense que la mejor estrategia sea la de requerir pruebas de manera secuencial. De esta manera ahorra costes por la utilización de pruebas que pueden no ser necesarias. Pero hay algo más.

Supongamos que tiene delante un paciente de esas características. Decide pedir una prueba para ver si tiene la enfermedad que usted sospecha. Esa prueba es positiva en el 90% de los enfermos (sensibilidad) y negativa en el 90% de los sanos (especificidad). Imaginemos que la prueba fue negativa. Si usted decide no seguir ese camino y buscar otra enfermedad que explique la clínica, se equivocará el 10% de las veces que aplique la prueba sobre un enfermo. Cuidado, no se equivocará el 10% de las veces que aplique la prueba. Repasemos estos conceptos.

Si aplicamos esta prueba a una población en la que 400 son enfermos y 600 son sanos. La probabilidad de enfermedad en cada uno de ellos es del 40%. Una probabilidad a priori que hace que el rendimiento de la prueba sea óptimo, como vimos en la entrega anterior. A medida que la probabilidad de enfermedad aumenta o disminuye, el rendimiento de la prueba disminuye. Entre los enfermos, el test habrá sido positivo en 360 y negativo en 40. Entre los sanos habrá sido positivo en 60 y negativo en 540. De manera que hay 40 falsos negativos del total de 580 negativos: el 94% de los negativos están sanos o el 6% están enfermos y no el 10% como se pudiera pensar. Si usted está dispuesto a asumir ese riesgo, no hará más pruebas. Pero si no lo está, porque la enfermedad es muy seria y el tratamiento muy beneficioso, habrá seguido una estrategia equivocada, porque el si el resultado hubiera sido positivo, el 86% (360 verdaderos positivos sobre 420 positivos) de los así clasificados están enfermos. Si no está dispuesto a que un 14% de los que va a tratar estén sanos, también necesitaría más pruebas para mejorar su especificidad.

En general, uno sigue investigando a los que resultaron positivos y deja de hacerlas a los negativos. Es el caso, por ejemplo, del cribado: sólo se hacen más estudios a los positivos a la mamografía. Con la siguientes pruebas se trata de disminuir la frecuencia de falsos positivos, en definitiva, incrementar la especificidad, como se puede ver en la tabla 1. Al coste, como puede adelantarse, de disminuir la sensibilidad; porque siempre que se incrementa una de estas capacidades de la prueba se erosiona la otra. Si usted no cree que se pueda tolerar perder ese 6% de pacientes, no debe empezar con una sola prueba. Tenga en cuenta que hemos puesto un ejemplo que va en contra del cribado: allí, con una incidencia de 1,5 por mil, el riesgo que asume de perder un caso es de 1,7 por 1000.

Supongamos que para esta enfermedad, además de ese test, que llamamos A disponemos de otro test al que llamamos B que tiene una sensibilidad del 80% y una especificidad del 60% ¿qué pasa si aplico los dos test a la vez en este paciente? Asumamos que las pruebas son independientes, es decir, que si es negativa en A no tiene por qué serlo en B. Esta es una suposición necesaria para el cálculo pero infrecuente en la práctica. Por ejemplo, en cribado de mama, la mamografía no es independiente de la ecografía. En estos casos, sólo estudios específicos de cómo funcionan las dos juntas nos permitirían hacer cálculos.

Cuando aplicamos dos pruebas a la vez, basta que una sea positiva para que se mantenga la sospecha, aunque sea falso positivo. Por eso, bajo la hipótesis de independencia, para ser un falso negativo tendría que serlo en A y en B a la vez. Es decir, el producto de esas dos probabilidades: $0.1 \cdot 0.2 = 0.02$. La sensibilidad, que es el complemento, será del 98%. La especificidad será la probabilidad de ser verdadero negativo en A (0.9) y en B (0.6), es decir, 0.54. Ha aumentado la sensibilidad al coste de tener más falsos positivos. De hecho, siguiendo esta estrategia, vamos a tener que estudiar con más pruebas a nada menos que a 276 sanos para hacer los 392 diagnósticos verdaderos. Si las pruebas que empleamos a partir de este momento no son muy específicas, aún trataremos a muchos sanos.

En la tabla 1 se muestra este ejemplo desarrollado con sus valores.

Tabla 1: Efecto de la utilización de pruebas, bien en serie o en paralelo, con el supuesto de prevalencia de enfermedad del 40%

Test	Sensibilidad	Especificidad	Valor predictivo positivo	Valor predictivo negativo
A	90	90	86	93
B	80	60	57	82
A y B en paralelo	98	54	59	96
A y B en serie	72	96	92	84

Se preguntará que cómo es posible que la estrategia de utilizar A y B en serie tenga el mismo resultado comoquiera que se utilicen, que no importa el orden. Esto es así sólo bajo la hipótesis de independencia de las pruebas. Recuerde que el orden de los factores no altera el producto: sensibilidad el producto de ambas ($0.9 \cdot 0.8 = 0.72$) y especificidad el complementario del producto de falsos positivos ($\{(1-0.4) \cdot 0.1 = 0.96\}$). Entreténgase usted en comprobarlo haciendo las tablas de dos por dos. Observará que si se empieza por el test A, como tiene mejor especificidad, menos pacientes serán examinados por B. Pero el resultado es que con ambas estrategias clasificará el mismo número de sujetos como positivos, 312 de los cuales 288 están sanos.

¿Cuándo emplear las pruebas en paralelo y cuándo en serie? Observe en la tabla que al final lo que comparamos son dos pruebas, pues la unión de A y B en serie resulta en una prueba con una determinada sensibilidad/especificidad y en paralelo en otra, como se ve en las filas 4 y 5 del cuadro. Dependerá, por tanto, del juego de valores con los que uno se tiene que manejar. Tenemos que darle un valor a la utilidad de tratar al sano y al enfermo y de no tratar al sano y al enfermo. Y conviene introducir en la utilidad el coste de la prueba, incluido el de realizar más consultas, en caso de estrategia en serie. La mejor forma de hacerlo es con un análisis de decisión, como se ha visto en la entrega anterior.

El clínico es un tomador de decisiones envuelto en la incertidumbre. Su mayor preocupación es no tratar, o tratar equivocadamente, a un paciente. Por eso prima siempre la sensibilidad de las pruebas: Prefiere la estrategia en paralelo. No tiene tanta preocupación o conciencia por los falsos positivos, por el daño que hace a clasificando a un sano, aunque sea sólo provisionalmente, como enfermo. Ni tampoco su primera preocupación es el coste que supone para el sistema este error. Sin embargo, los estudios que hay sobre la yatrogenia son alarmantes. Producimos salud, es cierto, pero también enfermedad. Será imposible, con los instrumentos que tenemos no cometer errores. Tenemos la obligación de minimizarlos, conociendo el alcance de nuestra tecnología y empleándola de la manera más adecuada a cada caso.

Palabras clave:

De la evidencia científica a las decisiones clínicas, Especificidad y sensibilidad de las pruebas diagnósticas

e-notas de evaluación

Autores/as: *Martín Caicoya Gómez-Morán*

Título: *Estrategia diagnóstica: ¿pescar con redes de arrastre o palangre?*

Nº 1 de 2006

Disponible en: *[Enlace a la noticia](#)*