

Notas metodológicas

De la evidencia científica a las decisiones clínicas

Las trampas en la causalidad

5/12/2006

Martín Caicoya Gómez-Morán

Jefe de Servicio de Prevención de Riesgos Laborales del Principado de Asturias. Hospital Monte Naranco.

Era el 24 de marzo de 1882. Entonces contagionistas y miasmáticos luchaban por la primacía en la explicación de tanta día ,Robert Koch reunió a lo más granado de la ciencia. En la sala había instalado su laboratorio. "Si la importancia de la para la humanidad se midiera por el número de muertes que causa, entonces la tuberculosis debe ser considerada importante que las más temidas enfermedades infecciosas, como la peste, el cólera y sus semejantes. Uno de cada humanos muere de tuberculosis. Si se considerara sólo a los grupos de edad productivos, la tuberculosis se lleva un tercio muchas veces más"¹. Al acabar su lectura se produjo un silencio. Ehrlich, un fisiólogo alemán que unos años más tarde recibió el premio Nobel, diría: "Mantengo que esa tarde viví la experiencia más importante de mi vida científica". No es de como los otros científicos, pudieron ver el bacilo que causa la tuberculosis con sus propios ojos a través de las lentes. Era el primer germen infeccioso demostrado. Este descubrimiento alumbró los famosos postulados de Henle-Koch para configurar el pensamiento causal hasta que los pulverizara la necesidad de explicar la relación entre tabaco y cáncer. Koch, y los científicos de su época, pensaban que la causa para ser tal debería ser necesaria (sin ella no hay enfermedad suficiente (si está presente, debe haber enfermedad).

Unos años antes, Gregory Mendel, tras un ordenado y minucioso estudio, iba a dictar sus hoy famosas leyes. Mendel presentó sus trabajos en las reuniones de la Sociedad de Historia Natural de Brno en 1865, publicándolos posteriormente como *Ergebnisse über die Bastardierung der Erbsen* sobre híbridos de plantas (*Versuche über Pflanzenhybriden*), en 1866, en las actas de la sociedad. Pero sus hallazgos fueron ignorados por completo. Los guisantes, esa leguminosa que adquiere un celebrado protagonismo en la mesa cuando llega primavera, fueron el fundamento de su éxito y la causa de su ostracismo.

Francis Galton, otro de los nietos del gran Erasmus Darwin, empleó también los guisantes para investigar la herencia, pero desde una perspectiva totalmente diferente. Ignoro si Galton conocía el trabajo de Mendel. Se sabe que su primo Charles Darwin, sí, pero no le prestó atención. Un día de 1875, Galton distribuyó paquetes de guisantes a cada uno de los siete amigos invitados a comer, quizá una menestra. En cada paquete los guisantes eran de tamaño semejante y a la vez cada uno de ellos de un tamaño diferente. Les pidió a estos *country gentleman* que los plantaran y le devolvieran la cosecha en la siguiente primavera. Galton puso en un gráfico el peso de los guisantes padres frente al de los hijos y observó que describían aproximadamente una línea recta, figura 1². Era la primera vez que se pensaba en la regresión y la correlación.

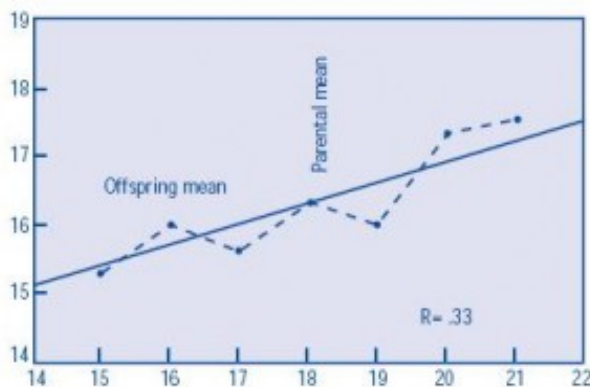


Figura 1

Figura 1: Gráfico de dispersión que muestra la relación entre el peso de los guisantes de los padres (eje X) y el peso de los guisantes de los hijos (eje Y). El eje X va de 14 a 22, y el eje Y va de 14 a 19. Se muestra una línea de regresión sólida que indica una correlación positiva. Una línea punteada representa la media parental, y otra línea punteada representa la media de la descendencia. El coeficiente de correlación R se indica como R = .33.

Pearson iba a perfeccionar matemáticamente estos dos descubrimientos, pero fue la necesidad que tenía Galton de encontrar una medida de asociación que respondiera a sus especulaciones hereditarias la que motivó el nacimiento de estas dos importantes herramientas estadísticas: la correlación y la regresión.

Galton tenía detrás toda la fuerza de la sociedad científica y abrió con sus experimentos la puerta de la investigación cuantitativa en genética que iba a florecer en ese país con aportaciones importantes a la estadística y la genética en los 50 años siguientes. Sin embargo, en vez de cavó un poco más la fosa donde estaba hundido el trabajo de Mendel. Se decía que sus leyes tal vez pudieran aplicarse a los guisantes, pero no a los seres llamados superiores.

Sin embargo, mientras en genética los postulados de Henle-Koch no tenían predicamento, en la ciencia médica eran incontestables hasta que la investigación cuantitativa en ciencias de la salud cobró auge en la segunda mitad del siglo XX. Fue el estadístico Bradford Hill, que precisamente con Richard Doll había demostrado una asociación entre tabaco y cáncer de pulmón que de una manera cumplía los postulados de Henle-Koch, quien por primera vez propuso unos criterios que contemplaran la multiplicidad de causas en un modelo que hoy empleamos. En la tabla 1 se muestran estos criterios³, tal como él los definió en 1965.

Tabla 1: Criterios de causalidad según Hill

FUERZA DE ASOCIACION	“La tasa de mortalidad por cáncer de pulmón”
CONSISTENCIA	“Se ha observado repetidamente por dife periodos” “La asociación se limita a trabajadores especo”
ESPECIFICIDAD	“De todas maneras, no debemos sobredimen” “Las enfermedades pueden tener más de un”
TEMPORALIDAD	“Es esa dieta la que produce la enfermedad que provocaron los cambios en los hábitos d”
GRADIENTE BIOLÓGICO	“O curva dosis-respuesta”
PLAUSIBILIDAD	“Lo que se considera biológicamente plausib”
COHERENCIA	“La interpretación de la relación causa-efecto los hechos conocidos de la historia natural y”
EXPERIMENTO	“Es posible, ocasionalmente, acudir a prueba”
ANALOGIA	“Con los efectos de la talidomida y la rubé aceptar pruebas menos contundentes, pero s en el embarazo.”

Obsérvese que no aparece como criterio el que la asociación sea estadísticamente significativa. Decía "No hay test formal de significación que puedan responder a estas preguntas. Esos tests pueden, y deben, recordarnos los efectos que el juego del azar puede crear, y nos instruirán en la posible magnitud de ese efecto, más allá de ahí, no contribuyen de ninguna manera a probar nuestra hipótesis".

En la revisión que Mervyn Susser ⁴ hizo de estos criterios considera que si cumple la significación estadística apoya la causalidad; y si no la cumple, siempre que el test tenga potencia, detrae. Pero lo más importante de su revisión es el concepto de "rendimiento predictivo". Porque eso esperamos de la ciencia: Queremos, que nos asegure que podremos modificar el curso natural de la enfermedad X (es decir, curarla) con el tratamiento. Predecir, y modificar, el futuro es nuestra aspiración. Por eso Susser le da a este criterio mucho valor, más que a la significación estadística.

El tiempo y la ciencia iban a dar la razón a Mendel. El no sabía que había genes (esta palabra la acuñó en 1909 el danés Johannsen) pero pensó que cada guisante tenía dos elementos hereditarios y que cada progenitor transmitía uno. También pensaba que uno de ellos era dominante, de manera que si los padres no transmitían el mismo elemento, el dominante daría el carácter. Las leyes de Mendel se cumplen fatalmente: recordemos la fenilcetonuria, o la enfermedad de Huntington, entre otros 2000 trastornos hereditarios descritos. En todos ellos se demuestra que un gen es necesario y suficiente para que se manifieste un determinado carácter.

También Galton tenía razón. Los biométricos decían, a los mendelianos, que era imposible que sus leyes se pudieran aplicar a los caracteres complejos porque no mostraban un patrón simple de herencia. Muchos son los caracteres que se heredan bajo la influencia de varios genes, de manera que uno no tiene, o raramente tiene, el 100% del carácter, tiene un porcentaje. De hecho, cuando son varios los genes que influyen en un carácter, los fenotipos que afectan se agrupan formando una distribución normal. Cuando uno está en el extremo derecho de la curva tiene mucha influencia de esa característica poligénica, por ejemplo, mucho colesterol. En otros casos, como la esquizofrenia, puede pensarse en un efecto umbral: a partir de un punto en la distribución se produce la sintomatología de la enfermedad. Se está o no se está enfermo. También puede pensarse que la enfermedad es gradual. De todas formas el medio modifica esas tendencias o predisposiciones.

La genética del comportamiento y la sociobiología son dos disciplinas muy interesantes que se basan en la genética cuantitativa. Especulan sobre la heredabilidad de ciertos rasgos de la personalidad, incluida la inteligencia que fue lo primero sobre lo que especuló Galton, padre de esta disciplina. Creo que uno de los mayores obstáculos en estas investigaciones es la dificultad de clasificar. Efectivamente, en la genética cualitativa la exposición está meridianamente clara: un gen tiene una o más bases diferentes, o simplemente tiene más bases, como pasa en el Huntington. La clasificación es bastante exacta y reproducible. Por otro lado, la enfermedad se define con precisión. No ocurre lo mismo con la genética cuantitativa, incluso con la técnica de análisis molecular llamada *quantitative trait loci* (QTL).

El progreso en el uso de esta técnica es importante, pero a pesar de ello, es difícil encontrar un gen que sea la fuente de la asociación hallada entre un carácter y un gen detectado por QTL. Se debe a que el margen de error, es decir, el intervalo de confianza al 90 % para la localización del QTL, es relativamente amplio ⁵.

Imaginemos que la sensibilidad de la prueba es del 80 % y la especificidad del 80 %. Vamos a pensar, para

facilitar el ejemplo, que los tests que se aplican para determinar un carácter psicológico clasifican bien, lo que pongo en duda. En la muestra, el 50 % tiene ese carácter psicológico, por ejemplo búsqueda de novedades. Las investigaciones tratan de asociarlo al receptor de dopamina DRD4. Supongamos que los que los expresen totalmente tengan 3,5 veces más probabilidades de manifestar esa característica de la personalidad. En la tabla 2 se muestra la realidad que hemos supuesto.

Tabla 2: Asociación (supuesta) entre DRD4 y búsqueda de novedad

PERSONALIDAD	DRD4 POSITIVO	DRD4 NEGATIVO	TOTAL
BUSQUEDA	60	40	100
NO BUSQUEDA	30	70	100
TOTAL	90	110	200

El odds ratio de esta prueba es $60 \cdot 70 / (40 \cdot 30) = 3,5$

Pero como la prueba de QTL no clasifica bien, de los 90 DRD4 positivos encuentra sólo 72 (sensibilidad de 0,8 multiplicado por los 90 que tienen la característica) y como la tasa de falso positivo es del 20% (1-especificidad) de los 110 que no tienen la característica, a 22 considera que la tienen (110 multiplicado por 0,2). A los 72 verdaderamente positivos hay que añadir los 22 falsos positivos: 94. Los DRD4 negativos serán 106. Ahora bien, esta operación hay que hacerla con las celdas. El resultado es el que figura en la tabla 3.

Tabla 3. Asociación encontrada entre DRD4 y búsqueda de novedad

PERSONALIDAD	DRD4 POSITIVO	DRD4 NEGATIVO	TOTAL
BUSQUEDA	56	44	100
NO BUSQUEDA	38	62	100
TOTAL	94	106	200

El OR es ahora de 2,1, se ha diluido la asociación por un problema de mala clasificación no diferencial porque la sensibilidad y especificidad aplicadas son idénticas para los que tienen la característica de personalidad y los que no la tienen. La mala clasificación no diferencial diluye las asociaciones, con lo cual hace fracasar la investigación. Es un problema extendido. En la práctica lo más frecuente es encontrar asociaciones de 2. En esos casos basta una pequeña mala clasificación para echar al traste todo el esfuerzo. Y es muchas veces causa de que unas veces sí y otras no se encuentre la asociación. Una falta de consistencia.

Un problema mucho más trascendente es la mala clasificación diferencial. Ocurre cuando el investigador emplea un instrumento clasificatorio para el grupo de sujetos con la característica y otro para el grupo de sujetos que no la tienen. Eso produce sesgos, asociaciones falsas. Contra eso hay que protegerse con un diseño adecuado. Pero no hay sitio aquí para tratarlo.

Decía Jeffrey Rose en una serie de artículos publicados hace años en BMJ con el título que yo he tomado prestado para esta serie, que la mala clasificación no diferencial era suciedad, ruido que impedía ver. Pero la mala clasificación diferencial era para él, y yo así lo creo, suciedad sucia. Ambas son enemigas del investigador, pero más la segunda.

Bibliografía

1. Tomado de <http://nobelprize.org/medicine/educational/tuberculosis/readmore.html> y traducido por el autor.
2. Stanton JM. Galton, Pearson, and the Peas: A Brief History of Linear Regression for Statistics Instructors. *Statistics Education*.2001;9.En <http://www.amstat.org/publications/jse/contents%5F2001.htm>.
3. Austin Bradford Hill, "The Environment and Disease: Association or Causation?," *Proceedings of the Royal Medicine*1965, 58, 295-300.
4. Susser M. Falsification, verification and causal inference in epidemiology.Reconsiderations in the light of Sir K philosophy. En *Epidemiology, health and society*. Oxford. Oxford Univeersity Press 1987.
5. Cabre JC. Complex behaviour and mouse models. En *Behavioral Genetics in the Postgenomic Era*. Plomin R, Craig IW, Mcguffin P. Washington. American Psychological Association 2002.

Palabras clave:

Causalidad, De la evidencia científica a las decisiones clínicas

e-notas de evaluación

Autores/as: *Martín Caicoya Gómez-Morán*

Título: *Las trampas en la causalidad*

Nº 8 de 2006

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