

Temas y controversias en bioestadística

Medwave 2014 Sep;14(8):e6016 doi: 10.5867/medwave.2014.08.6016

Causalidad y predicción: diferencias y puntos de contacto

Causality and prediction: differences and points of contact

Autor: Luis Carlos Silva Ayçaguer⁽¹⁾

Filiación:

⁽¹⁾Escuela Nacional de Salud Pública, La Habana, Cuba

E-mail: icsilva@infomed.sld.cu

Citación: Silva LC. Causality and prediction: differences and points of contact. *Medwave* 2014 Sep;14(8):e6016 doi: 10.5867/medwave.2014.08.6016

Fecha de envío: 24/7/2014

Fecha de aceptación: 4/9/2014

Fecha de publicación: 10/9/2014

Origen: solicitado

Tipo de revisión: revisado por pares

Resumen

La presente contribución expone las diferencias que existen entre aquellas variables que pudieran desempeñar un papel causal en determinado proceso y las que solo tienen un valor para pronosticar su desenlace. Se hacen algunas consideraciones sobre la medular intervención tanto de la asociación como de la precedencia temporal y los sesgos en ambos casos, el estudio de la causalidad y la elaboración de modelos predictivos. En ese contexto se examinan someramente varios aspectos relevantes relacionados con el diseño de los estudios correspondientes y se ilustran algunos de los errores que suelen cometerse al manejar tanto la causalidad como la predicción.

Abstract

This contribution presents the differences between those variables that might play a causal role in a certain process and those only valuable for predicting the outcome. Some considerations are made about the core intervention of the association and the temporal precedence and biases in both cases, the study of causality and predictive modeling. In that context, several relevant aspects related to the design of the corresponding studies are briefly reviewed and some of the mistakes that are often committed in handling both, causality and prediction, are illustrated.

Introducción

Es posible que a los lectores les resulten familiares expresiones del tipo "se ha encontrado asociación entre el consumo de X y la aparición de la enfermedad Y", "puede considerarse que el procedimiento X inhibe el desarrollo del proceso Y" o "el factor X resulta ser un buen predictor de la dolencia Y", las cuales suelen figurar en el apartado de Conclusiones de los artículos.

En un trabajo anterior de esta propia serie [1] se resaltó e ilustró que una conclusión como la primera de ellas, por muy extendido que esté su empleo, carece por lo general de interés. Aquilatar la asociación en sí misma entre dos variables no procede como objetivo de un estudio, aunque constatar su existencia sea crucial para poder hacer afirmaciones como la segunda o la tercera. Las preguntas de investigación jamás son estadísticas; la estadística es un mero intermediario metodológico para contestar preguntas sustantivas, sean sobre salud pública o sobre

un problema clínico o de las ciencias básicas. Consecuentemente, tampoco las respuestas pueden ser estadísticas. Y la estimación de una asociación es un asunto estadístico, no sustantivo.

Muchas veces el verdadero y genuino interés del investigador es demostrar (o por lo menos valorar) si X *influye* en Y, o si X permite de algún modo anticipar o *pronosticar* el valor de Y; pero como estas son "palabras mayores", el autor del informe se cura en salud quedándose en la constatación estadística de que los factores o variables analizadas exhiben asociación.

Es bien conocido que suele ser importante corroborar que la presunta causa (o factor de riesgo o de protección) esté asociada con la variable que registra el desenlace (típicamente, una enfermedad). La asociación es condición *sine qua non* de la causalidad y puede medirse a través de

diversos estadígrafos tales como, por ejemplo, un riesgo relativo, un odds ratio, un coeficiente de correlación o una diferencia de medias aritméticas (en este último caso, se trata de la asociación entre una variable cuantitativa y una dicotómica). Cuanto más intensa sea, más verosimilitud adquiere la hipótesis de que entre dichas variables se verifica una relación de causalidad. Ello se debe a que, tras identificar una asociación en los estudios observacionales, una de las más importantes incertidumbres estriba en que la correlación observada pudiera ser solo un reflejo del efecto que ejercen otros condicionantes, posibilidad que se minimiza cuando la asociación es "grande" [2].

¿Cuáles son las condiciones mínimas adicionales que han de cumplirse para seguir pensando que X gravita o influye sobre Y? ¿Cuál es la diferencia, en este contexto, entre el afán de estudiar causalidad y el de valorar la capacidad pronóstica de una condición dada? ¿También deben cumplirse otras condiciones en este último caso? Y de ser así, ¿son las mismas que las que exige la causalidad? Responder a estas preguntas es la finalidad del presente trabajo.

Condiciones para mantener en pie una hipótesis de causalidad

Las premisas (adicionales a la asociación) que han de verificarse para sostener como probable la conjetura de que cierta relación entre dos variables es causal, o que la presencia de un factor entraña un riesgo efectivo de enfermar (factor de riesgo), o que es un factor útil a los efectos de rehabilitarse o protegerse (factor de protección), se enumeran y comentan a continuación.

Aunque una hipótesis que proclama una relación causal ha de cumplir requisitos de entorno, tales como la plausibilidad biológica -condición que exige que la afirmación conjeturada cuente con cierto respaldo teórico y que no niegue verdades científicamente demostradas- aquellos metodológicamente cruciales, a los que insoslayablemente ha de atenderse, son los siguientes:

- **Congruencia temporal:** puesto que la presunta causa ha de haber actuado antes que el efecto, el diseño del estudio tiene que garantizar que las mediciones se hayan realizado con acuerdo a ello.
- **Ausencia de sesgos:** ocasionalmente los estudios padecen de defectos que conducen a que apreciemos de forma distorsionada en un sentido u otro la realidad; el inventario de sesgos posibles es cuantioso (de anticipación en el diagnóstico, del interrogador, de memoria, de selección, etcétera) y dependen del tipo de estudio que se realice. La literatura que los aborda es igualmente copiosa.
- **Carácter intrínseco de la asociación observada:** hay que corroborar que la asociación observada no puede ser enteramente atribuible a la acción de otros factores; es decir, ha de constatarse que la asociación subsiste una vez "controlados" los llamados factores de confusión, que en esencia son variables que influyen sobre las dos que se examinan de manera que puedan crear la ilusión de

que estas últimas guardan una relación esencial, no inducida por el efecto sobre ambas de terceras variables.

Cada una de estas tres premisas merece consideraciones independientes que desbordan los objetivos del presente artículo. Aquellos lectores interesados en profundizar sobre estos temas pueden acudir a las referencias correspondientes [3],[4],[5].

Diseños investigativos para el estudio de la causalidad

El tema de los diseños apropiados para investigar relaciones de causalidad en el marco salubrista y epidemiológico también exhibe un alto nivel de complejidad y ofrece muy diversas aristas que merecen atención detenida. Sin embargo, a los efectos que aquí interesan, basta con un recordatorio general sobre las alternativas básicas.

Los diseños concebidos para valorar hipótesis causales pueden ser tanto observacionales (de casos y controles o de cohorte) como experimentales. Estos últimos, de los cuales el ensayo clínico controlado es su expresión más emblemática en el mundo biomédico, constituyen la forma óptima para encarar el problema de la causalidad. El carácter prospectivo inherente a todo experimento resuelve el escollo posible de la temporalidad y, si el estudio se conduce adecuadamente -en particular cuando se usan técnicas de enmascaramiento y asignación aleatoria en el caso de los ensayos clínicos- muchos de los posibles sesgos resultan evitados por el propio diseño. Sin embargo, su virtud fundamental consiste en que, a través de la asignación aleatoria, especialmente si las muestras son grandes, los grupos que van a compararse se equiparan o balancean respecto de todos los factores que pudieran ser confusores, sean conocidos o no.

Lamentablemente, diversas realidades éticas y prácticas suelen imposibilitar la experimentación en el marco de la epidemiología y la salud pública. Los estudios observacionales de cohorte son los sucedáneos naturales del ensayo clínico, ya que comparten con él su naturaleza prospectiva: los hechos se registran en orden temporalmente ascendente. La observación comienza tanto para quienes ostentan una presunta condición de riesgo, como para quienes no la tienen y cuando aún no se han expresado los desenlaces que se estudian en uno y otro grupo. De ese modo, el registro adecuado de la temporalidad queda garantizado de antemano. Pero este enfoque alarga los lapsos para llegar a resultados analizables (a veces durante años) y suele ser prohibitivamente caro. Por eso suelen preferirse los estudios de casos y controles, a pesar de que ellos pueden ofrecer dificultades para asegurar el cumplimiento de la premisa de precedencia temporal [6].

Ahora bien, el gran problema que afecta a ambos tipos de estudios no experimentales lo ofrece el control de factores confusores. ¿Cómo asegurarse de que la asociación observada es intrínsecamente causal y que no se debe preponderante o exclusivamente a la influencia de al menos

un factor sobre los dos que se examinan? Se ha intentado dar respuesta a esta pregunta en cientos de artículos y libros. No procede aquí examinar en detalle tan delicado asunto, tanto por su complejidad como por el hecho de que no resulta necesario hacerlo para cumplimentar los propósitos de esta contribución. No obstante, cabe mencionar que el recurso clásico para ello era la llamada postestratificación y posterior aplicación de las técnicas propuestas por Mantel y Haenszel a finales de la década de los 50 del siglo pasado [7]. Sin embargo, tal procedimiento está en franco desuso debido a su naturaleza restringida y de mero valor exploratorio, y ha quedado, casi por entero, relegado a su nostálgica aparición en los cursos ortodoxos de epidemiología. Su lugar ha sido ocupado por los modelos multivariados, muy especialmente por la regresión logística [8].

Aunque los rituales unos generales, otros más puntuales para identificar factores de riesgo adoptan muchas formas y se multiplican, hay uno que parecería florecer especialmente. Lamentablemente, el poderoso instrumento de la regresión logística suele aplicarse con ese fin de manera ingenua y ceremonial. Se trata en esencia de un proceso que transita por los siguientes pasos:

1. Identificar un conjunto de variables que verosímilmente pudieran considerarse factores de riesgo y hacer un estudio "univariado" donde se valora la asociación existente entre cada uno de dichos factores con la variable de respuesta (típicamente una enfermedad) mediante una prueba de significación (por lo general Ji-Cuadrado o t de Student según se trate de una variable dicotómica o continua, respectivamente).
2. Ajustar un modelo logístico con las variables para las que se obtuvo "significación" en el paso anterior.
3. Aplicar un procedimiento algorítmico de selección de variables (regresión paso a paso) para determinar algo a lo que se le llama "modelo final", en el cual quedarían las variables o factores detectados, que son finalmente declarados como los verdaderamente influyentes.

La aparición de este proceso en artículos científicos, tesis y estudios epidemiológicos es tan frecuente que da la impresión de que muchos investigadores, a fuerza de su recurrencia, quedan ingenuamente más y más persuadidos de haber tropezado con una especie de "ábrete sésamo" con el cual pueden esclarecerse los misterios de la causalidad. No es difícil corroborar [9], [10] cuán estéril termina siendo su machacona aplicación cuando se emplea como una receta que prescinde de la reflexión crítica.

Predicción

Antes que nada, es preciso puntualizar qué se está entendiendo como "pronóstico" de determinado desenlace. Obviamente, un vaticinio inequívoco es imposible. Esto obliga a que toda predicción se tenga que verificar en un marco de incertidumbre. Consecuentemente, el manejo inicial siempre se produce en términos de la probabilidad que se calcula para los acontecimientos objeto de pronóstico, aunque operativamente, en un segundo momento, pueden aplicarse reglas que anuncian

determinado desenlace concreto sobre la base de dichas probabilidades. Por ejemplo, en el caso de una embarazada con cierto riesgo obstétrico, si la probabilidad de que se produzca un aborto supera cierto umbral, pongamos por caso, 0,8, se puede declarar a determinados efectos prácticos (y actuar en consecuencia) "que se va a producir" determinado evento salvo que se adopten tales y más cuáles medidas.

En este punto, debe advertirse algo sistemáticamente olvidado: aunque con mucha regularidad los conceptos de "factor de riesgo" y "factor predictivo" se emplean indistintamente, no son equivalentes. El hecho de que esté actuando o no un factor de riesgo (y conocer el grado en que influye) puede ocasionalmente ser útil para la predicción, pero una variable puede hacer importante aportación a los efectos de predecir, aunque en sí misma no sea un factor de riesgo ni tenga papel causal alguno.

Por ejemplo, la pérdida sostenida de peso puede tener valor predictivo a los efectos de que un adolescente realice un intento suicida, aunque tal pérdida no tenga como es obvio ninguna "responsabilidad" en el hecho. Y en esa calidad, ese dato puede tener interés a los efectos de la prevención. Toda variable explicativa tiene valor predictivo, pero lo que el ejemplo ilustra es que el recíproco es falso: un factor puede servir para la predicción sin que eso lo convierta en un agente causal.

¿Qué condiciones ha de cumplir una variable para contribuir sensiblemente a la formulación de un pronóstico?

Por una parte, al igual que ocurre en los estudios de causalidad, los sesgos constituyen un elemento perturbador que ha de ser evitado. Consideremos el más insidioso de ellos, el llamado "sesgo de selección", cuya eliminación, en los estudios de casos y controles, es absolutamente medular. Los procedimientos para elegir los controles (sujetos no enfermos) se pueden establecer adecuadamente siempre que se hayan fijado nítidamente los criterios empleados para seleccionar los casos (sujetos enfermos). El grupo control teóricamente más apropiado corresponde a la subpoblación de individuos que, habiendo estado en riesgo de padecer el desenlace objeto de estudio, en el hipotético caso de haber llegado a él, hayan podido ser elegidos como "casos". Es decir, ha de precaverse que ningún factor se halle más representado en uno de los grupos como consecuencia del modo en que ellos fueron conformados. Dicho de otro modo: un factor puede aparecer con más frecuencia entre casos que entre controles (lo cual haría pensar que tiene un papel causal), pero hay que descartar que tal diferencia haya sido inducida por la forma en que se seleccionaron los grupos. Si la condición de exposición influye de alguna manera en la posibilidad de que un individuo sea o no incluido como control, se producirá un sesgo de selección. Imaginemos que se valora el grado en que el consumo sistemático de carne cruda permite pronosticar la enfermedad de Alzheimer a partir de cierta edad. Supongamos que para estudiarlo se toman como "casos" todos los sujetos registrados como portadores de dicha enfermedad en el hospital de una región en la cual éste es

el único servicio que atiende tal dolencia, y como controles, a una muestra de pacientes libres de la enfermedad que arriban al servicio de urgencia de dicho hospital. En esta situación podríamos estar generando un sesgo de selección. En efecto, el consumo de carne cruda es más frecuente en zonas rurales que en el entorno urbano del hospital. Y ocurre que mientras el servicio que atiende el problema psiquiátrico recibe pacientes de toda la región, el servicio de urgencia atiende solo o muy predominantemente, casos urbanos (las urgencias de las zonas rurales acuden a centros de sus propios entornos). Siendo así, la muestra de controles tendría de antemano una subrepresentación de consumidores de carne cruda respecto de los casos, y ese monto adicional de consumidores de carne cruda entre los casos, que nada tiene que ver con la enfermedad mental, podría contribuir a que erróneamente le confirmamos valor para pronosticar.

Diseñar las preguntas de suerte que contemplen la debida precedencia temporal es asimismo fundamental para considerar que cierto factor pueda tener valor pronóstico para determinado desenlace.

Por ejemplo, consideremos la pavorosa epidemia de enfermedad renal crónica que desde hace poco más de 10 años está devastando algunas zonas de varios países (muy especialmente, algunas regiones de El Salvador, Nicaragua, Sri Lanka y la India). En estos enclaves, la enfermedad se expresa desproporcionadamente en jóvenes trabajadores agrícolas hasta el punto de producir tasas de prevalencia varias veces superior a lo que pudiera concebirse como normal [11]. Se ha considerado que la exposición a agroquímicos, tanto por contacto directo y prolongado, como por contaminación residual de larga data en el suelo, las aguas y los cultivos, pudiera tener responsabilidad causal en la epidemia [12].

Supongamos que nos preguntamos acerca del grado en que el hecho de ser un trabajador agrícola que labora en esas zonas permite vaticinar la aparición de la dolencia. En un estudio sobre esta temática [13] se pregunta a los sujetos de la muestra (sanos y enfermos) acerca de su exposición ocupacional, pero se indaga acerca de lo que éste hace actualmente (se registra su "ocupación actual"). Es decir, no se procura conocer en qué trabajaba el individuo antes de enfermar. Sin embargo, bien puede estar ocurriendo que algunos enfermos se dediquen actualmente a otra actividad económica (por ejemplo, a la venta ambulante), precisamente porque su estado le impide mantenerse en la actividad responsable de su dolencia. En tal caso, esta notable insuficiencia contribuirá a difuminar la asociación entre la exposición ocupacional y el trastorno, y por ende a ocultar en buena medida la posible capacidad de aquella para vaticinar su aparición.

Satisfechas estas dos demandas, ¿puede ser útil el recurso arriba descrito que concluye con la selección algorítmica de un modelo de regresión? La respuesta es afirmativa: cuando la técnica no se usa con finalidad explicativa sino pronóstica, parece enteramente razonable buscar el modelo más sencillo, y no hay razones para objetar el uso de rutinas algorítmicas (como la regresión paso a paso) para hallarlo. En ese caso es irrelevante que ocurran cosas

como que una variable que pudiera ser «directamente causal» resulte eliminada o quede suplida por una o más variables que no tengan influencia real alguna pero que se vinculen con aquella de manera que la predicción sea en definitiva eficiente, o por lo menos aceptable.

Desde el punto de vista del diseño, cabe apuntar que los tres procedimientos básicos (estudios experimentales, de cohortes o de casos y controles) pueden servir para hacer modelos de predicción. Puede sorprender que incluso estos últimos, siendo retrospectivos, puedan cumplir tal función. Sin embargo, incluso en ese caso es posible siempre que se cuente con cierta información complementaria, externa al estudio propiamente dicho. Por ejemplo, consideremos un estudio sobre factores de pronóstico de supervivencia de sujetos que han sufrido quemaduras y que arriban a un servicio especializado. Se toman sujetos que murieron (casos) e individuos que egresaron vivos del hospital (controles). Si se conoce la tasa global de mortalidad en dicho servicio de quemados, el modelo resultante puede corregirse y producir una función válida para pronosticar la sobrevivencia de un sujeto allí ingresado [14].

Finalmente, ha de señalarse que, a diferencia de lo que ocurre en los estudios de causalidad, cuando se examinan posibles variables predictoras, no es necesario descartar los factores confusores. Consideremos que se están construyendo tablas que permitan estimar el riesgo de muerte por concepto de enfermedad cardiovascular. Tal es el caso, por ejemplo, del llamado modelo SCORE (*Systematic Coronary Risk Evaluation*) [15]. Dicho modelo estima el riesgo de muerte en el curso del próximo decenio atendiendo a varias variables, tales como la edad, el sexo, la presión arterial sistólica, el colesterol total del suero y el tabaquismo actual. El tabaquismo se maneja dicotómicamente: fuma o no. Imaginemos que contamos con un instrumento que identifica el grado de pigmentación que tienen los dedos de cada sujeto y que, en lugar de preguntarle si fuma o no, simplemente se identifica si la pigmentación (derivada de sostener el cigarrillo en la mano) supera o no determinado umbral. Si en lugar de construir el indicador usando la variable tabaquismo se incorpora esta variable dicotómica basada en el grado de pigmentación, los resultados hubieran sido virtualmente los mismos. Obviamente, si un individuo fumador pudiera eliminar el color amarillo de sus dedos mediante algún tipo de jabón, ello no modificaría en absoluto su riesgo de morir en los próximos diez años debido a que dicha pigmentación no desempeña papel causal alguno en el proceso que se intenta vaticinar; pero a los efectos del vaticinio, esa naturaleza no causal es irrelevante, a la vez que resulta igualmente útil para calcular el riesgo.

Consideraciones finales

En numerosos artículos se anuncia que "el trabajo procura identificar factores de riesgo correspondientes a determinada dolencia" a pesar de que lo que realmente se hace es identificar factores que "se asocian" a dicho trastorno. Tal proceder pierde de vista que el hallazgo de una asociación es un paso que se queda a medio camino de lo que realmente se quiere. En otros casos, más elaborados, lo que efectivamente se hace es detectar variables con

valor predictivo, pero que no constituyen verdaderos factores de riesgo (no son factores que entrañen una relación de causalidad entre exposición y efecto) aunque el lenguaje empleado raramente lo reconoce, quizás, al menos en parte, porque el autor confunde esos conceptos y los da por equivalentes.

Tal equívoco se hace tangible, por ejemplo, cuando se emplean procedimientos tales como la construcción de modelos multivariados seguidos de la identificación de algo a lo que suele llamarse "modelo final", el cual es obtenido a través de un procedimiento algorítmico tal como la regresión paso a paso. Este procedimiento puede ser capaz de identificar variables pronósticas, pero jamás conseguirá separar los factores posiblemente causales de los que no lo son.

Notas

Declaración de conflictos de intereses

El autor ha completado el formulario de declaración de conflictos de intereses del ICMJE traducido al castellano por *Medwave*, y declara no haber tener conflictos de interés con la materia que aborda el artículo. El formulario puede ser solicitado contactando al autor.

Referencias

1. Silva LC. Manejo de la temporalidad en estudios observacionales para establecer relaciones causales. *Medwave* 2014;14(4):e5944 | [CrossRef](#) |
2. Hill AB. The environment and disease. Association or causation? *Proc R Soc Med*. 1965 May;58:295-300. | [PubMed](#) | [Link](#) |
3. Silva LC, Benavides A. Causalidad e inobservancia de la premisa de precedencia temporal. *Rev Methodol*. 1999;7:1-14. | [Link](#) |
4. Szklo MF, Nieto J. *Epidemiology: beyond the basics*. Gaithersburg: Aspen, 2000.
5. De Irala J, Martínez MA, Guillén F. ¿Qué es una variable de confusión? *Med Clin*. 2001 Oct;117(10):377-385. | [Link](#) |
6. Silva LC. Estudios de casos y controles en psiquiatría: causalidad, diseño y advertencias. *Actas Esp Psiquiatr*. 2004;32(4):236-248 | [Link](#) |
7. Mantel N, Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease. *J Natl Cancer Inst*. 1959 Apr;22(4):719-48. | [PubMed](#) |
8. Silva LC, Barroso I. *Regresión logística*. Madrid, España: La Muralla, 2003.
9. Silva LC, Barroso I. Selección algorítmica de modelos en las aplicaciones biomédicas de la regresión múltiple. *Med Clin*. 2001 May;116(119):741-745. | [Link](#) |
10. Silva LC. Una ceremonia estadística para identificar factores de riesgo. *Rev Salud Colectiva*. 2005;1(3):329-322. | [Link](#) |
11. Orantes, C.M. et al. Epidemiological characterization of chronic kidney disease in agricultural communities in El Salvador. *Medic Review* 2014;16(2):23-30. | [Link](#) |
12. Declaración de San Salvador. Abordaje integral de la enfermedad renal túbulo-intersticial crónica de Centroamérica (ERTCC) que afecta predominantemente a las comunidades agrícolas, 2013. salud.gob.sv [on line]. | [Link](#) |
13. Peraza S, Wesseling C, Aragon A, Leiva R, García-Trabanino RA, Torres C, et al. Decreased kidney function among agricultural workers in El Salvador. *Am J Kidney Dis*. 2012 Apr;59(4):531-40. | [CrossRef](#) | [PubMed](#) |
14. Silva LC. *Excursión a la regresión logística*. Madrid, España: Diaz de Santos, 1999.
15. Conroy RM, Pyörälä K, Fitzgerald AP, Sans S, Menotti A, De Backer G, et al. Estimation of ten-year risk of fatal cardiovascular disease in Europe: the SCORE project. *Eur Heart J*. 2003 Jun;24(11):987-1003. | [CrossRef](#) | [PubMed](#) |

Correspondencia a:

Calle 27e M y N #110
Vedado 10400
Ciudad de la Habana
Cuba



Esta obra de Medwave está bajo una licencia Creative Commons Atribución-No Comercial 3.0 Unported. Esta licencia permite el uso, distribución y reproducción del artículo en cualquier medio, siempre y cuando se otorgue el crédito correspondiente al autor del artículo y al medio en que se publica, en este caso, Medwave.