Revisión a los clásicos.

Conocer, clasificar, ordenar

24/06/2011

Martín Caicoya Gómez-Morán

Coordinador de Programas del Servicio de Calidad de la Dirección General de Calidad e Innovación de la Consejería de Salud y Servicios Sanitarios

El fenómeno de Will Rogers phenomenon. Migración de estadios y nuevas técnicas diagnósticas como una fuente de error en las estadísticas de supervivencia de cáncer

Feinstein AR, Sosin DM, Wells CK. N Eng J Med 1985. 312 : 1604-8.

Al comparar la supervivencia, tanto global como ajustada por estadio, de la cohorte de pacientes con cáncer de pulmón tratada en 1977 con la de 1953-1964, se observó que la primera era más larga. Pero a la cohorte más reciente se le había realizado muchas pruebas de imagen nuevas. En este trabajo se realizó el ejercicio de estadiar a la cohorte de 1977 con los métodos de 1952-64. El resultado es que se evidencia un sesgo de adelantamiento que hace que su pronóstico sea más favorable. Además, se descubre que con las nuevas técnicas se encontraban metástasis que antes no podían ser identificadas, hay una migración de estadios. Pacientes que antes se hubieran clasificado en un "buen estadio" fueron asignados a un "mal estadio". El resultado es que la supervivencia por estadio es mejor en 1977 pero cuando se estadia con el método de 1953-64 no hay diferencias. De hecho, si la clasificación se realiza por sintomatología, que no está influida por las nuevas técnicas, las dos cohortes tienen supervivencia similar.

Comentario

La naturaleza está ordenada. Esta fue la premisa básica para acceder al saber. Y saber, desde esa perspectiva, es simplemente la descripción de la naturaleza. Pero no puede ser una descripción como un relato, donde las cosas suceden unas detrás de las otras. Conocer la naturaleza es clasificarla, porque orden y clasificación son dependientes. Sin embargo, la enfermedad es una manifestación de desorden. Si no hay orden no hay clasificación. Así pensaban Hipócrates y Aristóteles. Fue Sydenhan1 el primero que de una manera sistemática se impuso la tarea de clasificar las enfermedades: para patologías específicas habrá remedios específicos. Perdura en ciertos ámbitos esta denominación para los medicamentos. Sydenhan pretendía una taxonomía de la enfermedad de la misma manera que los botánicos clasificaban las plantas.

Pero su tarea era mucho más difícil. Le faltaba una teoría que pudiera relacionar entre sí los datos de sus observaciones.

Aunque hoy la medicina que se llamó empírica, porque se basaba en la forma de conocer que recomendaba Bacon, no tiene apenas seguidores, el esfuerzo clasificatorio sigue siendo fundamental. Está en la base de la epidemiología, junto con contar y comparar. Hay todo un cuerpo de estudio y conocimiento que tiene que ver con los defectos de clasificación, bien diferenciales o no diferenciales. Los primeros se denominan sesgos.

Precisamente a la búsqueda y demostración de sesgos dedicó parte de su carrera Alvan Feinstein. Es conocido por la demostración del sesgo de detección en el diagnóstico de cáncer de útero en mujeres que toman estrógenos en la menopausia2 y el de cáncer de pulmón en fumadores3 . Pero quizá su contribución más ingeniosa ha sido el que el llamó, originalmente, fenómeno de Will Rogers.

Will Rogers fue un vaquero que trabajó en circos y rodeos y más tarde se hizo un nombre como cómico en cines y teatros. Entre sus chistes había uno que era parecido a esto: "Cuando los de Oklahoma emigraron a Califorinia se elevó el nivel de inteligencia en ambos estados". No hace falta decir que Rogers era de Oklahoma.

Una tentación para evaluar la eficacia de los tratamientos es comparar la supervivencia entre diferentes periodos. Pero los controles históricos son malos compañeros porque hay muchas otras cosas que influyen en el éxito del tratamiento, además de las nuevas tecnologías. Entre otras cosas, el momento, en el curso natural de la enfermedad, en que se realiza la intervención. Influye de dos maneras: por un lado en que si se diagnostica antes, naturalmente vivirá más; por otro porque es posible que el tratamiento precoz sea más eficaz. Lo primero lo conocemos como sesgo de adelantamiento, que también demuestra Feinstein en esta publicación. Lo segundo, en cáncer, se pretende controlar comparando supervivencia ajustada por estadio.

La originalidad del trabajo que aquí comento de Alvan Feinstein es que se dio cuenta de que lo mismo que los más tontos de Oklahoma emigran a California, país donde todos son aún más tontos, cuando se estadia con métodos más precisos, se produce una migración de los casos más avanzados al estadio superior. Eso quiere decir que los que quedan son posiblemente de mejor pronóstico y los que inmigran del estadio inferior no solo no agravan su situación sino que la mejoran. Por tanto es fácil encontrar una mejor supervivencia estadio específico a través de los tiempos simplemente por un fenómeno clasificatorio. Debido a ello no se debe comparar supervivencia con controles históricos incluso ajustando por estadio.

Además hay otra razón, no argumentada por Feinstein, que en mi opinión es más importante. A medida que aumenta la cobertura y prestaciones sanitarias y paralelamente se eleva la conciencia de salud de la población, más diagnósticos precoces se realizan. Muchos de estos cánceres quizá nunca hubieran progresado o lo hubieran hecho muy lentamente. Aquí no hay un problema de estadios, hay un problema de biología del tumor: los cánceres estadio I de próstata o mama tienen mucho mejor pronóstico en 2011que en 1980, antes de la era del cribado. Es lo que se denomina sesgo de duración o lenght bias4 .

Volviendo al trabajo original de Feinstein, los métodos de clasificación empleados en cada uno de los periodos de estudio, 1953-64 y 1977, fueron diferentes. Lo que ocurrió es que mejoró la validez del método, es decir, su sensibilidad y especificidad. Por eso se produce un error de clasificación. Es un error diferencial porque se emplean métodos diferentes en función del grupo de comparación. Si lo que hubiera ocurrido fuera que cada uno de los métodos se empleara al azar, en los dos grupos, el error de clasificación sería no diferencial. Un error no diferencial sería el que se cometiera por el temblor del pulso con una escopeta con el punto de mira bien calibrado. Un error diferencial ocurre si la escopeta está mal calibrada. En una clasificación con error no diferencial distribuido al azar la media es la medición correcta: la diana. En el diferencial la media es el error de calibración. Imaginemos que se hubiera evaluado la eficacia de un tratamiento frente a otro para el estadio I en esos dos periodos históricos. En la cohorte antigua había un 22% de pacientes en ese estadio y en la moderna un 18%. Parecería que no hay adelantamiento del diagnóstico. La supervivencia a los 6 meses fue en 1953-64 de 75% y en 1972 del 92%, diferencias significativas. Pero cuando se clasifica la cohorte nueva con el método de 1953-64 ocurre que se incrementa el número de pacientes en estadio I, precisamente porque con las nuevas técnicas muchos habían migrado a estadios superiores. Ahora ya son 36%. Además de un sesgo de clasificación hay un sesgo de adelantamiento que no se detectaba. Con esta clasificación no hay diferencias en la supervivencia: 75% frente a 76%. Esto produce un dilema: ¿Se puede aplicar la efectividad de los estudios históricos cuando se realizaron con una clasificación menos precisa del TNM? Lógicamente, los beneficios serán mayores en la actualidad, pero quizá también se perjudique a pacientes con buen pronóstico.

Un problema semejante de clasificación lo sufrió el colesterol en las primeras investigaciones del asturiano Grande-Covián5. Entonces solo se medía rutinariamente el colesterol total. Es bien sabido que demostraron que los ácidos grasos saturados de cadena media, en concreto C14 a C18, elevan el colesterol, los poliinsaturados lo rebajan y los monoinsaturdos no lo modifican. Eso puso en duda la bondad del aceite de oliva. Es también un error de clasificación, pero su perfil es algo diferente.

Bibliografia:

1- Sydenham, T;: Locke, J. Ensayo Sobre La Anatomía. KRK editores. 2009.

2- Horwitz R I, Feinstein A R, Horwitz S M, Robboy S J. <Previous Article|Next Article> Necropsy diagnosis of endometrial cancer and detection-bias in case/control studies. Lancet 1981;8237: 66-68

3- McFarlane M J, Feinstein A R, Wells C K Necropsy evidence of detection bias in the diagnosis of lung cancer. Arch Int Med 1986; 146: 1695-1698

4- Shwartz M. Estimates of Lead Time and Length Bias in a Breast Cancer Screening Program. Cancer 46:844-851, 1980

5- Keys, A; Anderson, JT y Grande, F. Serum cholesterol response to changes in the diet. The effect of cholesterol in the diet. Metabolism. 1965 14: 759-765

Palabras clave:
will rogers phenomenon
métodos de clasificar

Autores:
Martín Caicoya Gómez-Morán

Nº:26 de 2011