



Comparación de resultados quirúrgicos: un mínimo control, por favor

Sr. Director:

Nadie negará que esté de moda comparar resultados quirúrgicos entre centros o servicios, incluso dentro de un mismo departamento, entre cirujanos especializados y los no especializados. El artículo de Luján et al¹ es un ejemplo de lo segundo, y mucho me temo que sea también un clamoroso ejemplo de excesiva simplificación en cuanto al método. Se simplifica en exceso cuando se compara tasas o proporciones brutas, sin control, entiéndase, sin ajuste o estandarización. El hecho de comparar tasas brutas de 2 grupos (cirujanos expertos en una materia y cirujanos generales) mediante una avalancha de pruebas estadísticas clásicas, la *t* de Student, Kaplan-Meier con Logrank o Breslow, la χ^2 , etc., no añade ajuste alguno en las comparaciones, con lo cual, casi seguro, los resultados estarán sesgados, aunque a la postre pudieran apuntar en la dirección verdadera. Creer lo contrario es un error de concepto bastante común en esta clase de artículos, y no sólo en literatura quirúrgica en español.

Esto ocurre porque en comparaciones no aleatorizadas de resultados siempre se entrometen factores de confusión e interacción, es decir, variables que por su propia

cuenta influyen en el resultado, a veces tanto o más que la variable principal que se compara, en este caso, el tipo de cirujano. Por supuesto, una buena aleatorización resolvería el problema, pero aplicarla en esta clase de estudios es imposible la mayoría de las veces, por razones prácticas y por razones puramente éticas. Métodos de ajuste hay muchos y no se trata de describirlos ahora, pero me gustaría explicar lo que se podría considerar un mínimo exigible. Se trata de lo que se conoce en epidemiología como ajuste indirecto² de riesgos, que utiliza la propia muestra o población de enfermos de cuyos datos disponemos en un momento dado, a falta de unos estándares generales o universales, que en la materia que tratamos no suelen estar confeccionados, o por lo menos, a mí no me consta que existan.

Imaginemos que se quiera comparar el resultado de los 2 grupos de cirujanos, expertos y no expertos, en cuanto a recidivas locales en el cáncer de recto. Básicamente, procederíamos de la siguiente forma con nuestra tabla de datos:

Paso 1. Calcular el riesgo simple y general de recidiva local (G): número total de recidivas acontecidas dividido por el número total de enfermos analizados. Se trata de una proporción en bruto de ambos grupos de cirujanos en conjunto.

Paso 2. Calcular el riesgo de recidiva de cada paciente, sin contar aún con qué clase de cirujano lo operó, pero ajustando por las variables de influencia que puedan intervenir: rasgos del tumor, tratamientos previos, rasgos de los pacientes relacionados con la dificultad quirúrgica, etc. Esto se debe conseguir mediante regresión logística en la que la variable de respuesta sea la presencia o ausencia de recidiva local y todos los posibles factores de influencia entrasen como variables predictoras. Debo insistir en que aquí todavía no debe entrar en consideración la variable cirujano.

Paso 3. Como en el paso anterior, al programa estadístico que utilicemos le habremos pedido que guarde la probabilidad de recidiva de cada paciente según el modelo logístico; entonces podremos sumar esas probabilidades individuales para los pacientes tratados por los especialistas por un lado, y las de los no especialistas, por otro. Esa suma es el número esperado (E) de recidivas para cada grupo de cirujanos, de acuerdo con los factores independientes del cirujano que pudieran influir en el resultado de dichas recidivas locales.

Paso 4. Sumemos ahora las recidivas reales que ha tenido cada grupo de cirujanos por separado. Es el número observado (O) de recidivas locales.

Paso 5. Calculemos el cociente O/E de cada grupo, que debemos multiplicar por el riesgo general de recidivas (G). Así obtenemos finalmente la estimación ajustada del riesgo de recidiva local para cada tipo de cirujano.

Estas 2 cifras de riesgo ajustadas son ya comparables directamente, sin necesidad de pruebas de hipótesis, aunque también las haya para ello. Por otra parte, si los esperados son más que los observados, se tratará de un buen resultado individual para un grupo de cirujanos, y si ocurre al contrario, el resultado de un grupo, previo a

cualquier comparación, será... malo. Igualmente, si el riesgo así ajustado supera el 1, también es un mal resultado, mientras que si es menor de 1, el resultado es bueno para el grupo. Naturalmente, se precisa un juicio de valor para catalogar las diferencias de riesgo ajustado entre ambos grupos como grandes o pequeñas, o como tolerables o intolerables. Pero lo importante es saber que aquí no gana necesariamente el que menos recidivas recolecta en bruto, ni siquiera cuando el valor p es significativo. Tener eso claro es el mínimo que pido por favor.

Javier Escrig Sos

Servicio de Cirugía. Hospital General de Castellón.
Castellón. España.

Bibliografía

1. Luján J, Hernández Q, Valero G, De las Heras M, Gil J, Frutos D, et al. Influencia del factor cirujano en el tratamiento quirúrgico del cáncer de recto con radioquimioterapia preoperatorio. Estudio comparativo. Cir Esp. 2006;79:89-94.
2. Colimon KM. Fundamentos de epidemiología. Madrid: Díaz de Santos; 1990.