

Qué son y cómo se interpretan los *outcomes* compuestos

Gabriel Rada^{1,2,3}, Macarena Roa^{1a}.

How to interpret composite outcomes

El *outcome* es aquello que se mide en un estudio. El resultado es el valor numérico que adquiere el *outcome*. Si un estudio evalúa la disminución de la mortalidad (*outcome*), el resultado es, por ejemplo, que la mortalidad disminuye en 10%. En la traducción al castellano de la palabra *outcome* se han utilizado muchos sinónimos: evento, desenlace, observable, variable de resultado, etc. Considerando que ninguno de ellos goza de popularidad o consenso, hemos preferido mantener el término en inglés.

Un *outcome* compuesto (*composite outcome*) es aquel que incorpora más de un elemento. Se contabiliza como positivo cada vez que un paciente cumple con alguno de ellos.

Por ejemplo, en el estudio ACCORD se evaluó el efecto de un tratamiento hipoglucemiante intensivo en la prevención de eventos cardiovasculares, para lo cual se definió un *outcome* compuesto (OC) de accidente vascular encefálico, infarto agudo al miocardio y muerte cardiovascular¹. Cada vez que un paciente presentaba alguno de estos eventos, se contabilizaba que presentó el *outcome*.

A lo largo de este artículo utilizaremos preferentemente como ejemplo el estudio ACCORD, que ha sido analizado críticamente en el presente número de la *Revista Médica de Chile*.

POR QUÉ SE UTILIZAN OC

Existen dos razones principales para utilizar OC en los estudios clínicos:

1. La primera es de sentido común: ya que generalmente existe más de un posible beneficio de una intervención, es razonable medir si cualquiera de éstos se produce, es decir, el efecto neto de la intervención. Siguiendo nuestro ejemplo, se podría suponer que si la intervención X disminuye los accidentes vasculares es tan importante como si disminuye los infartos miocárdicos.
2. La segunda es una razón estadística; para poder demostrar su hipótesis, el investigador necesita que su estudio incluya un determinado número de pacientes (tamaño muestral), de manera de poder tener confianza en que los resultados que se obtendrán no se deben al azar. Uno de los factores de los que depende el tamaño muestral es del número de veces en que se espera que ocurra el *outcome* (riesgo basal o tasa de eventos en el grupo control), siendo menor el tamaño muestral requerido mientras más frecuente sea el *outcome*. Al utilizar un OC, se suman los eventos que aporta cada uno de los componentes. Por esta razón, mientras más componentes decida agre-

¹Unidad de Medicina Basada en Evidencia, Pontificia Universidad Católica de Chile.

²Departamento de Medicina Interna, Pontificia Universidad Católica de Chile.

³Unidad Docente Asociada, Hospital Dr. Sótero del Río. Santiago de Chile.

^aResidente

Correspondencia a: Dr. Gabriel Rada. Módulo Universidad Católica, Hospital Dr. Sótero del Río, Av. Concha y Toro 3459. Fax: 56-2-2897638. E mail: gabriel@rada.cl

gar un investigador, menor será el tamaño muestral que requiera.

En un artículo de esta serie se ha revisado con mayor profundidad el cálculo de tamaño muestral².

¿POR QUÉ EL USO DE OC HA IDO EN AUMENTO EN LOS ÚLTIMOS AÑOS?

Esta práctica ha aumentado con un ritmo diferente en distintas áreas de la investigación en medicina, siendo en la de estudios cardiovasculares donde se ha observado el mayor crecimiento en el uso de OC³. Una revisión de los estudios clínicos randomizados publicados en 14 de las más importantes revistas en cardiología, constató que 73% utilizaban un OC como *outcome* primario, y que prácticamente todos utilizaban algún OC⁴. Otros estudios muestran algo similar en otras áreas⁵.

Una de las principales razones para explicar este aumento se relaciona con el aspecto estadístico discutido anteriormente: a medida que el riesgo basal ha ido disminuyendo (por la mejoría de las condiciones de vida, por los avances de la medicina o porque se comienzan a estudiar pacientes en etapas iniciales de la enfermedad), y que los efectos de las terapias son menores (porque se adicionan a otros tratamientos, o simplemente porque hoy en día se evalúan tratamientos menos efectivos), los investigadores necesitan de un mayor número de pacientes o de un mayor tiempo de seguimiento para poder demostrar sus hipótesis.

A modo de ejemplo, a principio de la década 1980-89 el infarto agudo al miocardio con supradesnivel del segmento ST tenía 20% de mortalidad en nuestro país⁶. En la actualidad, se ha logrado reducir a 10%⁷. Si un investigador quisiera determinar la efectividad de un tratamiento que teóricamente disminuye a la mitad la mortalidad, en los años 80 hubiese necesitado alrededor de 400 pacientes. Actualmente, necesitaría cerca de 900. Si hacemos el cálculo considerando que el efecto de la intervención fuera una reducción de 20% en vez de 50% (una cifra bastante más realista), se necesitaría alcanzar un tamaño muestral sobre 6.000 pacientes.

Así, con el uso de OC, pueden realizarse estudios que de otro modo serían poco factibles y muy costosos. Probablemente esta situación se irá agudizando en el futuro; por un lado la medicina

seguirá experimentando progresos, y por otra parte, a medida que se avanza en el entendimiento de las enfermedades y se realizan más estudios, el «nicho» al que van apuntando los investigadores va cambiando. Para ilustrar lo anterior, a fines de la década de 1980 varios estudios demostraron que los inhibidores de la enzima convertidora de angiotensina disminuyen la mortalidad en pacientes con insuficiencia cardíaca postinfarto agudo al miocardio^{8,9}. Los estudios posteriores evaluaron si también lo hacían en pacientes con infarto pero sin insuficiencia cardíaca¹⁰⁻¹³. El paso siguiente fue analizar pacientes sin infarto pero con alto riesgo coronario¹⁴, y luego se realizaron estudios en pacientes con un riesgo cardiovascular menor¹⁵. A medida que se estudian pacientes con cada vez menos riesgo basal, el tamaño muestral necesario va aumentando. Los investigadores tienen 3 alternativas: conseguir los recursos para alcanzar el tamaño muestral requerido, aumentar el tiempo de seguimiento (y así permitir que ocurran más eventos) o utilizar OC.

Precauciones al utilizar outcomes compuestos en la toma de decisiones. Si bien la utilización de OC puede tener ventajas tanto para los que toman decisiones clínicas como para los que desarrollan la investigación, es necesario que se cumplan algunos requisitos¹⁶. Si cualquiera de los puntos siguientes no se cumple, se corre un alto riesgo de que estas decisiones se vuelvan difíciles o imposibles.

1. La importancia de los distintos componentes es similar. No perdamos de vista que el resultado de los estudios lo utilizaremos finalmente para tomar decisiones clínicas. Es por eso que debemos tener en cuenta la importancia que tiene para el paciente cada uno de los elementos del OC, y las diferencias o gradiente de importancia entre ellos.

Al momento de informar sobre los efectos de un tratamiento, lo intuitivo es interpretarlo como uno solo. Si el compuesto contiene *outcomes* que son importantes para el paciente y otros que no lo son tanto, la interpretación y aplicación se vuelve más difícil. Por ejemplo, muchos estudios cardiovasculares incluyen *outcomes* compuestos por eventos como necesidad de revascularización, infarto agudo al miocardio o muerte cardiovascular⁴. Si le decimos a un paciente que determinado tratamiento disminuye la probabilidad de revascularizarse (la mayoría por vía endovascular), infartar-

se o morir, probablemente para el paciente es más importante evitar la muerte que evitar una revascularización. Si resulta que este tratamiento disminuye esta última, pero no las muertes, al utilizar la información de este OC, se induce a error.

En el caso del estudio ACCORD, se midió un *outcome* compuesto por infarto agudo al miocardio no fatal, accidente cerebrovascular no fatal o muerte por causas cardiovasculares. Los 3 componentes son importantes para el paciente, y si bien siempre existirá algo de gradiente, en este caso es suficientemente pequeña, por lo que podemos considerarnos satisfechos con respecto a este punto¹.

Una revisión de todos los estudios clínicos randomizados publicados en 6 de las revistas médicas más importantes, y que incluyeron OC, mostró que 54% de ellos tenían una gradiente moderada o alta entre sus componentes. Adicionalmente, los *outcomes* de menor importancia para los pacientes fueron los que ocurrieron con mayor frecuencia y en los que el tratamiento tuvo más efecto⁵.

2. Los componentes del outcome ocurren con igual frecuencia. Para que el OC sea informativo, los distintos componentes deben ocurrir con una frecuencia similar. Uno de los casos más recurrentes en que se produce desbalance es en los compuestos que incluyen el *outcome* mortalidad. Si bien se trata de un elemento de la mayor importancia, en muchas ocasiones tiene una frecuencia de ocurrencia mucho menor que el resto de los componentes.

En el estudio ACCORD la frecuencia con que ocurren los 3 componentes es bastante similar, por lo que podemos decir que se cumple adecuadamente este punto (infarto miocárdico 3,6% vs 4,6%; AVE 1,3% vs 1,2%; muerte de causa cardiovascular 2,6% vs 1,8%).

Es importante destacar que para poder evaluar este punto y el siguiente, es fundamental que los investigadores reporten los resultados de cada componente por separado, lo cual no ocurre en un tercio de los casos⁵.

3. La intervención tiene el mismo efecto relativo sobre los componentes del outcome en la teoría y en la práctica. La racionalidad de generar OC supone que la intervención en estudio tiene un efecto similar sobre cada uno de los componentes.

Es importante volver a recalcar que en la toma de decisiones clínicas necesitamos decir: «el tratamiento

tiene efecto sobre A, sobre B y sobre C» y por consiguiente se lo recomendaremos a los pacientes en que queremos afectar A, B y C. Ahora consideremos que en un paciente deseamos disminuir A y B, pero C no nos interesa (o incluso podemos querer evitarlo). O podemos pensar que C tiene causas distintas que A y B. ¿Y si el tratamiento disminuye A, pero aumenta B y C?, ¿cómo decidimos? ¿Qué pasa si aumenta A y B, pero sobre C no podemos estar seguros?

Para afirmar que los componentes de un *outcome* tienen el mismo efecto relativo en la teoría y en la práctica, se debe cumplir lo siguiente:

- La fisiopatología es similar.
- El riesgo relativo es similar, y con un intervalo de confianza estrecho.

Ilustremos lo anterior con algunos ejemplos:

En un estudio que evaluó el efecto del tratamiento anticoagulante oral por tiempo prolongado en pacientes con enfermedad tromboembólica, se reporta que el tratamiento anticoagulante disminuyó en 48% el *outcome* compuesto de recurrencia de tromboembolismo, hemorragia mayor y muerte, ¿qué entendemos en este caso? ¿El tratamiento anticoagulante disminuye el riesgo de tromboembolismo y también el de hemorragia? Si disminuye las muertes, ¿es por su efecto en el tromboembolismo o en las hemorragias? Como podemos observar, si bien la intención de este OC es entregar un «efecto neto» de la intervención, la información que entrega no es útil, ya que la fisiopatología que subyace tras los distintos componentes es muy distinta¹⁷.

El OC del ACCORD (infarto agudo al miocardio no fatal, accidente cerebrovascular no fatal o muerte por causas cardiovasculares) es muy razonable desde el punto de vista fisiopatológico, ya que los 3 componentes comparten los mismos mecanismos patogénicos y serían modulados por la intervención (control intensivo de la glicemia) de una forma similar.

Además de la racionalidad fisiopatológica, debemos fijarnos en si esta «similitud teórica» se cumple en la práctica. Esto lo hacemos observando los resultados individuales de cada componente. Siguiendo con el ACCORD, se observó una reducción significativa del riesgo relativo de infarto agudo al miocardio no fatal de 22% (intervalo de confianza 95%: 5 a 38), un incremento no significativo del riesgo de accidente cerebrovascular de 8% (intervalo de confianza 95%: -28 a 44) y un aumento significativo de las muertes por causas cardiovasculares de 44% (intervalo de confian-

za 95%: 13 a 76). Como podemos constatar muy claramente en este caso, el resultado individual de cada componente es completamente diferente: Uno beneficioso, uno neutro y uno perjudicial. Utilizar el resultado del OC (reducción de 10%, con intervalo de confianza 95% de -4 a 21) sin duda que conduciría a una toma de decisiones inapropiada.

CONCLUSIONES

La utilización de OC puede tener ventajas tanto para los que toman decisiones clínicas como para

los que desarrollan la investigación. Su uso ha ido aumentando en los últimos años y es muy posible que lo siga haciendo aún más. Al momento de interpretar esta forma de presentar los resultados, debemos respondernos 3 preguntas: ¿Existe una gradiente de importancia entre los componentes del *outcome*?, ¿la frecuencia de ocurrencia de los distintos componentes es la misma?, y ¿el efecto de la intervención sobre los distintos componentes es el mismo? Si cualquiera de estos puntos no se cumple satisfactoriamente, la toma de decisión en base a OC se hace difícil o imposible.

REFERENCIAS

1. ACTION TO CONTROL CARDIOVASCULAR RISK IN DIABETES STUDY GROUP. GERSTEIN HC, MILLER ME, BYINGTON RP, GOFF DC JR, BIGGER JT, BUSE JB ET AL. Effects of intensive glucose lowering in type 2 diabetes. *N Engl J Med* 2008; 358: 2545-59.
2. CANDIA R, RADA G. Estudios randomizados interrumpidos precozmente por beneficio: ¿Muy buenos o muy malos? *Rev Méd Chile* 2006; 134: 1470-5.
3. FREEMANTLE N, CALVERT M, WOOD J, EASTAUGH J, GRIFFIN C. Composite outcomes in randomized trials: greater precision but with greater uncertainty? *JAMA* 2003; 289: 2554-9.
4. LIM E, BROWN A, HELMY A, MUSSA C, ALTMAN D. Composite Outcomes in Cardiovascular Research: A Survey of Randomized Trials. *Ann Intern Med* 2008; 49: 612-7.
5. FERREIRA-GONZÁLEZ I, BUSSE JW, HEELS-ANSELL D, MONTORI VM, AKL EA, BRYANT DM ET AL. Problems with use of composite end points in cardiovascular trials: systematic review of randomised controlled trials. *BMJ* 2007; 334: 786.
6. ALMENDARES C, CORBALÁN R, CASTRO P, RODRÍGUEZ A, MARCHANT E, KUNSTMANN S ET AL. Hospital course of acute myocardial infarction: significance of the therapeutic procedures of early reperfusion. *Rev Méd Chile* 1995; 123: 1365-71.
7. PRIETO JC, SANHUEZA C, MARTÍNEZ N, NAZZAL C, CORBALÁN R, CAVADA G ET AL; GRUPO DE ESTUDIO MULTICÉNTRICO DEL INFARTO. Mortalidad intrahospitalaria en hombres y mujeres según terapias de reperusión en infarto agudo del miocardio con supradesnivel del ST. *Rev Méd Chile* 2008; 136: 143-50.
8. PFEFFER MA, LAMAS GA, VAUGHAN DE, PARISI AF, BRAUNWALD E. Effect of captopril on progressive ventricular dilatation after anterior myocardial infarction. *N Engl J Med* 1988; 319: 80-6.
9. SHARPE N, MURPHY J, SMITH H, HANNAN S. Treatment of patients with symptomless left ventricular dysfunction after myocardial infarction. *Lancet* 1988; 1: 255-9.
10. GRUPPO ITALIANO PER LO STUDIO DELLA SOPRAVVIVENZA NELL'INFARTO MIocardico. GISSI-3: effects of lisinopril and transdermal glyceryl trinitrate singly and together on 6-week mortality and ventricular function after acute myocardial infarction. *Lancet* 1994; 343: 1115-22.
11. ISIS COLLABORATIVE GROUP. ISIS-4: randomized study of oral captopril in over 50,000 patients with suspected acute myocardial infarction. *Circulation* 1993; 88: Suppl I: 1-394.
12. SWEDBERG K, HELD P, KJIEKSHUS J, RASMUSSEN K, RYDÉN L, WEDEL H. Effects of the early administration of enalapril on mortality in patients with acute myocardial infarction: results of the Cooperative New Scandinavian Enalapril Survival Study II (CONSENSUS II). *N Engl J Med* 1992; 327: 678-84.
13. AMBROSIONI E, BORCHI C, MAGNANI B. The effect of the angiotensin-converting-enzyme inhibitor zofenopril on mortality and morbidity after anterior myocardial infarction. The Survival of Myocardial Infarction Long-Term Evaluation (SMILE) Study Investigators. *N Engl J Med* 1995; 332: 80-5.
14. YUSUF S, SLEIGHT P, POGUE J, BOSCH J, DAVIES R, DAGENAIS G. Effects of an angiotensin-converting-enzyme inhibitor, ramipril, on cardiovascular events in high-risk patients. The Heart Outcomes Prevention Evaluation Study Investigators. *N Engl J Med* 2000; 342: 145-53.
15. BRAUNWALD E, DOMANSKI MJ, FOWLER SE, GELLER NL, GERSH BJ, HSIA J ET AL; PEACE TRIAL INVESTIGATORS. Angiotensin-converting-enzyme inhibition in stable coronary artery disease. *N Engl J Med* 2004; 351: 2058-68.
16. MONTORI VM, PERMANYER-MIRALDA G, FERREIRA-GONZÁLEZ I, BUSSE JW, PACHECO-HUERGO V, BRYANT D ET AL. Validity of composite end points in clinical trials. *BMJ* 2005; 330: 594-6.
17. RIDKER PM, GOLDBERGER SZ, DANIELSON E, ROSENBERG Y, EBY CS, DEITCHER SR, ET AL; PREVENT INVESTIGATORS. Center for Cardiovascular Disease Prevention and the Division of Preventive Medicine, Brigham and Women's Hospital and Long-term, low-intensity warfarin therapy for the prevention of recurrent venous thromboembolism. *N Engl J Med* 2003; 348: 1425-34.