

¿Qué son los puntos finales intermedios (surrogate outcomes)?. Parte 7

What are surrogate outcomes?

Bottaro FJ ^{1,2}

1. Servicio de Clínica Médica - Hospital Británico
2. Coordinador del Comité de Revisión Institucional - Hospital Británico

febottaro@hotmail.com

Fecha de recepción: 18/08/2015
Fecha de aprobación: 24/08/2015



COMO LEER
UN ARTÍCULO
CIENTÍFICO

HEMATOLOGÍA: 172 - 177
Volumen 19 n° 2
Mayo - Agosto 2015

Palabras clave: desenlaces intermedios, desenlace surrogado, punto final.

Keywords: surrogate outcomes, surrogate end point.

*“No existen problemas agotados, solo hay hombres agotados por los problemas.”
Luis Federico Leloir*

En un escenario ideal en la práctica clínica cotidiana, la mayoría de nuestras decisiones estarían basadas en los resultados de ensayos clínicos controlados que demuestren el efecto de una intervención determinada sobre un **desenlace o punto final** (en inglés: *outcome*). En este escenario, lo ideal sería que estos desenlaces representen en forma directa eventos que fueran *importantes para los pacientes* (DIP). Por ejemplo, modificación de la expectativa de vida (disminución de la mortalidad), reducción de complicaciones (disminución de eventos como infarto agudo de miocardio, accidente cerebrovascular, amputaciones, recidivas o recaídas de enfermedades tumorales) o modificación de la calidad de vida. Sin embargo, en muchas situaciones hallaremos que los resultados de un estudio se resumen en *desenlaces o puntos finales “intermedios”* (en in-

glés *surrogate outcomes*). Los desenlaces intermedios (DIM) suelen ser habitualmente mediciones de laboratorio, o variables fisiológicas que se utilizan en reemplazo o como sustitutos de los desenlaces importantes para los pacientes.

La sustitución de DIM por los DIP es muy atractiva cuando el desenlace intermedio puede ser medido en forma muy temprana, más fácilmente, más frecuentemente, con mayor precisión, o con menor riesgo de ser influido por otras terapias o factores de riesgo.

¿Son necesarios los desenlaces intermedios?

El uso de DIM permite realizar estudios de manera más rápida que con el uso de DIP, ya que se podría adjudicar antes la eficacia de una intervención sin tener que aguardar años hasta la ocurrencia de un evento determinado. Generalmente los DIM permi-

ten realizar estudios con menor tamaño muestral, más baratos, en menor tiempo, y muchas veces de manera más sencilla. Por estas razones podremos observar que muchos de los estudios que podemos hallar en la literatura médica se basan en DIM para concluir sobre los efectos de una intervención, especialmente los estudios financiados por la industria farmacéutica. Esto puede ser beneficioso y adecuado en algunos contextos en los que abreviar tiempos hasta la aprobación de una droga puede ser beneficioso en el tratamiento de algunas enfermedades. Por ejemplo, la agencia regulatoria de medicamentos norteamericana (FDA, Food and Drug Administration) ha basado la aprobación de muchas drogas antiretrovirales para el tratamiento de la infección por el virus de la inmunodeficiencia humana (HIV) en DIM. Muchas de estas drogas que fueron aprobadas bajo esta modalidad han demostrado su utilidad mediante su impacto en DIP (reducción asociada del síndrome de inmunodeficiencia adquirida y su mortalidad asociada) en estudios de cohortes prospectivos.⁽¹⁾

Sin embargo, la aprobación de drogas para su uso

comercial basado en su efecto en DIM puede ser perjudicial y resultar en aumentos de morbilidad y/o mortalidad. Existen múltiples ejemplos en la literatura de situaciones en las que ha ocurrido esto. Los casos de la milrinona, vesnarinona y xamoterol son ejemplos en los que demostrada eficacia en la mejoría de DIM en pacientes ambulatorios con insuficiencia cardíaca no tuvo su correspondiente correlato en DIP en ensayos clínicos controlados aleatorizados (ECCAs) que evidenciaron un exceso en la mortalidad en cada una de estas drogas.^(2, 3, 4, 5) En la literatura existen múltiples ejemplos y situaciones en las que una droga que demostró eficacia basada en DIM no ha podido mantener su aparente beneficio en DIP, o incluso luego se ha demostrado un rol nocivo. El lector interesado en leer información adicional al respecto y hallar varios ejemplos descriptos adecuadamente puede dirigirse a un interesante capítulo del libro de Guyatt: *User's Guide to the Medical Literature*.⁽⁶⁾ En la **tabla 1** se puede observar la diferencia entre DIP y DIM vinculados a similares contextos clínicos.

Tabla 1: Desenlaces intermedios y su desenlace clínicamente relevante

Niveles de colesterol LDL	Eventos cardiovasculares
Recuento de linfocitos CD4 en HIV	Infecciones vinculadas al SIDA
Niveles de presión arterial	Accidente cerebrovascular
Trombosis venosa profunda detectada por venografía	Enfermedad tromboembólica venosa clínica
Disminución niveles de glucemia	Complicaciones vinculadas a la diabetes

¿Qué condiciones debe cumplir un DIM para ser considerado un desenlace útil?

Para considerar útiles en la toma de decisiones en salud a los DIM, estos deben poseer una asociación fuerte e independiente con un DIP. Esto significa que el DIM debe ser predictor de eventos clínicos relevantes y todos los efectos de una intervención sobre un beneficio clínico deben ser “captados” por el DIM. Esto implica que no solo el mecanismo biológico o fisiopatológico del efecto de una intervención debiera ser conocido, si no también vía de acción.

Cuando el beneficio de un DIM no se traduce en un beneficio “importante para el paciente”, estamos frente a un desenlace irrelevante. Esta asociación entre DIM y beneficios clínicos relevantes no siem-

pre es lineal y lógica, y muchas veces nos encontramos que los ECCA con DIP no demuestran los beneficios observados en estudios previos con DIM. En la **tabla 2** se enumeran diferentes contextos clínicos en los que el uso de DIM resultó ser problemático. La literatura médica abunda de casos en que aún una excelente hipótesis fisiopatológica que podría ser comprobada mediante el uso de DIM puede ser perjudicial de acuerdo a lo que se observa en ECCA. Un ejemplo paradigmático es el estudio CAST⁽⁷⁾ (Cardiac Arrhythmias Suppression Trial). Este ensayo se basó en que las arritmias están asociadas a muerte súbita y son la principal causa de muerte en pacientes con infarto agudo de miocardio (IAM). Existen drogas anti-arrítmicas que pueden suprimir la actividad cardíaca eléctrica anormal.

Tabla 2: uso problemático de desenlaces intermedios

• Aumento de la densidad mineral ósea en osteoporosis sin traducción clínica de disminución del riesgo de fracturas
• Disminución de arritmias ventriculares posterior a infarto de miocardio, sin evidencias de mejoría en la sobrevida.
• Disminución de niveles plasmáticos de colesterol sin impacto en la incidencia de eventos cardiovasculares o mortalidad cardiovascular
• Disminución de niveles plasmáticos de hemoglobina glicosilada en diabetes sin impacto en la incidencia de complicaciones de la enfermedad
• Aumento de la contractilidad en insuficiencia cardíaca sin mejoría de la sobrevida
• Disminución de trombosis venosa profunda detectada por venografía sin evidencias claras de reducción de mortalidad y/o tromboembolismo de pulmón.

Por lo que los investigadores se plantearon que la administración de drogas anti-arrítmicas posterior a un infarto agudo de miocardio podría reducir muerte súbita y mortalidad total en pacientes pos-IAM. El estudio CAST aleatorizó (estudio doble ciego) cerca de 1500 pacientes a dos grupos de tratamiento (encainida y flecainida) y otro placebo. Se evidenció un exceso de mortalidad en ambos grupos tratamiento comparado con placebo.

Otro ejemplo interesante es el uso de fluoruro de sodio en mujeres posmenopáusicas para la prevención de fracturas osteoporóticas. Es reconocido que la pérdida de densidad mineral ósea (DMO) es progresiva en mujeres pos-menopáusicas. El fluoruro de sodio ha demostrado estimular la formación ósea e incrementar hasta un doble la masa ósea. Por estas observaciones se planteó que la administración de fluoruro de sodio podría reducir la incidencia de fracturas vertebrales en esta población de mujeres. Con este fin se realizó un ECCA que aleatorizó 202 mujeres pos-menopáusicas a recibir en forma doble ciega fluoruro de sodio versus placebo. Todas las mujeres fueron suplementadas con calcio. Llamativamente, y muy a pesar que el fluoruro de sodio evidenció un incremento de la DMO tanto en columna vertebral como en cuello de fémur (35%, $p=0,0002$ y 12%, $p=0,0002$ respectivamente), tanto las fracturas vertebrales como las no vertebrales fueron más frecuentes en el grupo tratamiento.⁽⁸⁾

¿Por qué hay que tener especial cuidado con estudios que basan sus resultados en DIM?

El uso inapropiado de DIM puede no solo a asumir efectos falsos de ciertas intervenciones, sino también a enmascarar u ocultar información acerca de potenciales efectos adversos de las misma. La capacidad de un ensayo clínico para detectar diferencias estadísticamente significativas en la tasa de eventos entre dos o más grupos depende del número de pacientes incluidos (tamaño de la muestra). Por lo que para observar eventos que ocurren con relativa baja frecuencia se requiere de estudios de grandes proporciones. Esta situación es frecuente en la literatura médica, y son reconocidos los ejemplos en los que recién luego de varios estudios clínicos se advierte la asociación de una intervención con un evento adverso indeseado que genera el retiro de la droga del mercado y la suspensión del desarrollo de la misma. Un ejemplo de esta situación es el de la rosiglitazona, un fármaco aprobado por la FDA en 1999 luego de varios estudios que demostraban que la droga reducía los niveles de glucemia y hemoglobina glicosilada plasmática.⁽⁹⁾ No fue hasta el año 2007 que se acumularon alertas del potencial riesgo cardiotoxicó de la rosiglitazona y su potencial asociación con insuficiencia cardíaca y el desarrollo de edemas de miembros inferiores.⁽¹⁰⁾

En ciertos contextos puede ser importante tener información sobre el efecto de una droga o interven-

ción en forma muy temprana; por ejemplo, la FDA basó las primeras aprobaciones de anti-retrovirales en estudios que se basaban en DIM. Sin embargo, en la mayoría de los contextos clínicos de las enfermedades crónicas prevalentes, es extremadamente importante que las intervenciones en evaluación no solo hayan demostrado eficacia en DIM, sino también en desenlaces importantes para los pacientes. También es valorable que exista evidencia de grandes cohortes prospectivas que no alerten sobre posibles eventos adversos inadvertidos en la etapa de investigación clínica y demuestren la asociación en la vida real de la intervención y desenlaces importantes para los pacientes.

El reconocimiento del alcance de los desenlaces intermedios es importante para interpretar en forma adecuada los resultados de estudios clínicos y para la adecuada toma de decisiones diariamente.

Un caso de análisis en las implicancias que posee el uso de DIM en lugar de DIP es el de la profilaxis de tromboembólica venosa (ETEV) en paciente quirúrgico. La mayoría de los ECCAs que avalan el uso de drogas anti-trombóticas para la prevención de ETEV peri-operatoria han utilizado como desenlace para medir el efecto de estos métodos de diagnóstico como la venografía aún en el paciente asintomático. Por lo que, el rastreo sistemático de trombosis venosa profunda asintomático en estos pacientes imposibilita que estos estudios posean importante tasa de eventos clínicos para poseer poder estadístico que demuestre efecto sobre eventos importantes para los pacientes (tromboembolismo o muerte por embolia de pulmón). Esta situación no sería problemática si la comunidad médica internacional considerase a la ETEV asintomática detectada en forma venográfica (DIM) como un predictor de ETEV clínica.⁽¹¹⁾ Sin embargo, y a pesar que es muy posible que este desenlace intermedio (ETEV venográfica) sea válido, es difícil estimar el impacto clínico neto de la reducción de eventos venográficos.⁽¹²⁾ Este ejemplo nos demuestra como a pesar que un desenlace impresione poseer un correlato clínico directo y pueda ser irrefutable, su validez puede ser cuestionada, y sin estudios apropiados tampoco se puede estimar el beneficio clínico que se desprendería de él. Este tipo de situaciones plantea un importante desafío no solo a nivel individual en la práctica diaria, sino también a nivel de las recomendaciones de sociedades científicas. Por ejemplo, las recomendaciones

del American College of Chest Physician (ACCP) del año 2008 distinguían con una recomendación clase **1A** el uso de heparinas de bajo peso molecular (HBPM), fondaparinux y antagonistas de la vitamina K (AVK) para la prevención de ETEV en cirugía ortopédica mayor programada. Llamativamente, estas mismas guías emitían una recomendación en contra del uso de aspirina en este contexto con un grado también **1A**.⁽¹³⁾ En la **tabla 3** se indica el significado de una recomendación **grado 1A**. Estas guías fueron realizadas basando sus recomendaciones en los desenlaces primarios de los ECCAs, siendo estos desenlaces intermedios (ETEV venográfica). Para el año 2012, con el fin de homogeneizar diferentes guías de práctica clínica y responder a múltiples críticas de la comunidad médica internacional^(14, 15, 16), el ACCP publicó su 9° consenso sobre conductas anti-trombóticas basando sus recomendaciones esta vez en eventos clínicamente importantes o desenlaces importantes para los pacientes.⁽¹⁷⁾ En el caso de la cirugía ortopédica mayor programada, el noveno consenso sugiere de preferencia el uso de HBPM por sobre otras drogas como fondaparinux, aspirina, AVK, heparina no fraccionada, dabigatrán, rivaroxaban y apixaban con un grado de recomendación **2B** (excepto para los AVK, 2C). La razón por la que se disminuye el nivel de evidencia es debido a que la ETEV venográfica aporta información indirecta como DIM de eventos clínicos realmente importantes, y porque los estudios no poseen poder suficiente para demostrar diferencias estadísticas en estos últimos. Es interesante destacar como desde los mismos estudios pueden surgir recomendaciones diferentes en base a los desenlaces que se tomen como parámetro para basar la recomendación, y como esto puede tener un impacto diferente en la práctica cotidiana.

Conclusiones

En esta entrega hemos repasado diferentes significado de los desenlaces intermedios o “surrogate outcomes” en contraposición o desenlaces que importan a los pacientes. Comprender como se desarrollan los estudios de investigación y como se interpreta la información que se desprende de ellos es importante para asimilar los resultados de los estudios a la toma de decisiones diaria. La toma de decisiones basada en estudios que solo miden DIM sobre una intervención determinada no necesariamente nos llevará a un beneficio clínico real, sino que además potencial-

mente puede ocultar eventos adversos relevantes y eventualmente ser perjudicial para nuestros pacientes. Los DIM además generan un importante proble-

ma en la valoración del equilibrio beneficio-riesgo por parte de nuestros pacientes.

Tabla 3: Grado de recomendaciones American College of Chest Physicians

Recomendación fuerte, alto grado de calidad de evidencia: Grado 1A	Beneficios claramente superiores a riesgos potenciales (o viceversa)	Evidencia consistente de EECAs sin limitaciones importantes, o excepcionalmente evidencia proveniente de estudios observacionales de excelente calidad	La recomendación puede ser aplicada a la mayoría de los pacientes en la mayoría de las circunstancias. Mayor evidencia es poco probable que cambie la confianza en los resultados
Recomendación debil, moderada calidad de evidencia: Grado 2B	Beneficios estrechamente equilibrados con riesgos	Evidencia de ECCA con importantes limitaciones (resultados inconsistentes, sesgos metodológicos, información imprecisa o indirecta) o evidencia muy fuerte de estudios observacionales	La mejor acción decisión puede diferir de acuerdo a circunstancias particulares del paciente o valores de la sociedad. Investigación de excelente calidad adicional podría tener un impacto importante en la confianza sobre los estimados y modificar la recomendación.

Bibliografía

- Olsen CH, Gatell J, Ledergerber B, y col. Risk of AIDS and Death at Given HIV-RNA and CD4 Cell Count, in Relation to Specific Antiretroviral Drugs in the Regimen. *AIDS*. 2005; 19 (3): 319-330.
- Timmis AD, Smyth P, Jewitt DE. Milrinone in heart failure. Effects on exercise haemodynamics during short term treatment. *British Heart Journal*. 1985; 54(1):42-47.
- Packer M, Carver JR, Rodeheffer RJ, col. Effect of oral milrinone on mortality in severe chronic heart failure. The PROMISE Study Research Group. *N Engl J Med*. 1991; 325(21):1468-75.
- Asanoi H, Sasayama S, Iuchi K, col. Acute hemodynamic effects of a new inotropic agent (OPC-8212) in patients with congestive heart failure. *J Am Coll Cardiol*. 1987;9(4):865-71.
- Cohn JN, Goldstein SO, Greenberg BH, y col. A dose-dependent increase in mortality with vesnarinone among patients with severe heart failure. Vesnarinone Trial Investigators. *N Engl J Med*. 1998;339(25):1810-6.
- Brignardello-Petersen R, Ioannidis P, Tomlinson G y col. Surprising results of randomized trials, en Guyatt G, Drummond R, Meade M, Cook D. *User’s Guide to the Medical Literature*. JAMAevidence. McGraw Hill Education. 2014.
- Echt DS, Liebson PR, Mitchell LB, y col. Mortality and morbidity in patients receiving encainide, flecainide, or placebo. The Cardiac Arrhythmia Suppression Trial. *N Engl J Med*. 1991;324(12):781-8.
- Riggs BL, Hodgson SF, O’Fallon WM, y col. Effect of fluoride treatment on the fracture rate in postmenopausal women with osteoporosis. *N Engl J Med*. 1990;322(12):802-9.
- Nolan JJ, Jones NP, Patwardhan R, col. Rosiglitazone taken once daily provides effective glycaemic control in patients with Type 2 diabetes mellitus. *Diabet Med*. 2000;17(4):287-94.
- Nissen SE, Wolski K. Effect of rosiglitazone on the risk of myocardial infarction and death from cardiovascular causes. *N Engl J Med*. 2007;356(24):2457-71.
- Chan NC, Stehouwer A, Hirsh J, y col. Lack of consistency in the relationship between asymptomatic DVT detected by venography and symptomatic VTE in thromboprophylaxis trials. *Thromb Haemost*. 2015; 114(5).
- Kassaï B, Shah NR, Leizorovicza A, col. The true treatment benefit is unpredictable in clinical trials using surrogate outcome measured with diagnostic tests. *J Clin Epidemiol*. 2005; 58(10):1042-51.

¿QUÉ SON LOS PUNTOS FINALES INTERMEDIOS (SURROGATE OUTCOMES)?

13. Geerts WH, Bergqvist D, Pineo GF, col; American College of Chest Physicians. Prevention of venous thromboembolism: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines (8th Edition). *Chest*. 2008; 133(6 Suppl):381S-453S.
14. Stewart DW, Freshour JE. Aspirin for the prophylaxis of venous thromboembolic events in orthopedic surgery patients: a comparison of the AAOS and ACCP guidelines with review of the evidence. *Ann Pharmacother*. 2013; 47(1):63-74.
15. Lachiewicz PF. Comparison of ACCP and AAOS guidelines for VTE prophylaxis after total hip and total knee arthroplasty. *Orthopedics*. 2009;32(12 Suppl):74-8.
16. Eikelboom JW, Karthikeyan G, Fagel N y col. American Association of Orthopedic Surgeons and American College of Chest Physicians guidelines for venous thromboembolism prevention in hip and knee arthroplasty differ: what are the implications for clinicians and patients? *Chest*. 2009; 135(2):513-20.
17. Falck-Ytter Y, Francis CW, Johanson NA, y col; American College of Chest Physicians. Prevention of VTE in orthopedic surgery patients: Antithrombotic Therapy and Prevention of Thrombosis, 9th ed: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines. *Chest*. 2012; 141(2 Suppl):e278S-325S.