

## ESTUDIOS DE NO INFERIORIDAD

CONSTANZA ROJAS B.<sup>1</sup> y DAVID TORRES P.<sup>2</sup>

- Los estudios con control activo pueden ser de superioridad (intervención > control), equivalencia (intervención similar al control) y no inferioridad (intervención “no es peor” que el control).
- Los estudios de no inferioridad son apropiados cuando se quiere evaluar la eficacia de un tratamiento experimental *versus* un control activo, cuando nuestra hipótesis es que el tratamiento experimental no es necesariamente más efectivo que el tratamiento probado, pero es estadísticamente “no inferior” (no es inaceptablemente menos efectivo).
- Los principales problemas de este tipo de estudios son: sensibilidad del ensayo, análisis según intención de tratar *vs* por protocolo, escoger el margen M, tamaño muestral. *Bio-creep*.
- Muchos estudios de este tipo no cumplen con los criterios necesarios para tener verdadera validez, aumentando la falsa aceptación de tratamientos inferiores como terapia alternativa, por lo que son altamente cuestionados.

La mayoría de los estudios que evalúan la eficacia de los tratamientos médicos son estudios clínicos aleatorizados controlados (ECA) con placebo o con control activo. Este tipo de estudios son más eficaces y requieren de un tamaño muestral relativamente pequeño para detectar la efectividad de un tratamiento. Cuando existe un tratamiento demostradamente efectivo, existen consideraciones éticas que impedirían entregar placebo a un grupo que sabemos que se va a beneficiar del tratamiento ya conocido, sólo para demostrar que el nuevo tratamiento también es útil. Por ejemplo, en la profilaxis de trombosis venosa profunda (TVP) en cirugía de cadera, el tratamiento con enoxaparina es conside-

rado efectivo y seguro. Existe un nuevo tratamiento (dabigatran), que pretende ser igual de efectivo en la prevención de TVP y con algunos beneficios en la adherencia al tratamiento. En este caso, no es ético realizar un ECA controlado con placebo, ya que existe un tratamiento efectivo para esta condición y el riesgo de no dar profilaxis a esta cirugía puede acarrear consecuencias potencialmente letales.

Esta situación no es infrecuente, ya que existe un gran número de tratamientos considerados efectivos para diversas condiciones. Es aquí donde entran en juego los estudios con control activo, los que se categorizan en estudios de superioridad (intervención > control), equivalencia (intervención similar al control) y no inferioridad (intervención ‘no es peor’ que el control). La mayoría de los médicos están familiarizados con el diseño y sesgos que se pueden encontrar en los ensayos clínicos controlados por placebo, sin embargo, ¿somos capaces de realizar un análisis crítico de un estudio de no inferioridad?

### ESTUDIOS DE NO INFERIORIDAD

Este tipo de estudio es apropiado cuando se quiere evaluar la eficacia de un tratamiento experimental *versus* un control activo (tratamiento probado), cuando nuestra hipótesis es que el tratamiento experimental no es necesariamente más efectivo que el tratamiento probado, pero es estadísticamente ‘no inferior’ (es decir, no es inaceptablemente menos efectivo).

Sin embargo, para poder establecer la no inferioridad de un tratamiento se deben cumplir óptimamente con 3 criterios:

- 1) El tratamiento a considerar no debe mostrar inferioridad terapéutica en relación al control activo.
- 2) El tratamiento exhibiría eficacia terapéutica si se realizara un estudio controlado con placebo.

<sup>1</sup> Perioperative Research Fellow, Population Health Research Institute, McMaster U. Canada.

<sup>2</sup> Clínica Santa María. Instructor, Universidad de los Andes.

3) El tratamiento ofrece ventajas secundarias en conveniencia, costo (ej. costo entre una angioplastia vs bypass coronario para tratamiento de angina), seguridad (ej. menos hemorragia con el uso de aspirina vs anticoagulantes posterior a un infarto) o tolerancia (anticoagulante oral vs inyectable para prevención de enfermedad tromboembólica).

Todo esto se debe cumplir en el contexto de una población de referencia adecuada, un control activo con dosis probada por ensayos clínicos controlados por placebo bien diseñados, un alto nivel de adherencia al tratamiento y un poder estadístico adecuado.

Al tener en cuenta todas estas especulaciones, requerimientos y conjeturas inherentes a un estudio de no inferioridad, es difícil no pensar en los sesgos y problemas que pueden venir asociados a este tipo de estudio. Es por eso que el clínico debe tener en cuenta un par de puntos claves que ayudarán a establecer la validez de un estudio de no inferioridad.

## DISEÑO

Al diseñar un estudio de no inferioridad hay que formular la hipótesis sobre la base de que el objetivo primario es probar que el efecto del tratamiento experimental (E) no es inferior (o menor o menos efectivo) que el efecto del control activo (C) dentro de un margen de no inferioridad o rango ( $\Delta$ ). En base a esto planteamos la hipótesis nula y la hipótesis alternativa de la siguiente forma:

$$H_0 = \text{EfectoE} - \text{EfectoC} > \Delta \quad H_A = \text{EfectoE} - \text{EfectoC} \leq \Delta$$

si se rechaza la hipótesis nula ( $H_0$ ) en relación a que el efecto experimental es inferior al efecto del control activo, el efecto del tratamiento experimental podría ser superior, equivalente o levemente inferior al control activo, sin embargo, lo que se puede asegurar es que la inferioridad no es mayor al margen  $\Delta$ .

Un ejemplo es la comparación de la eficacia del ximelagatran (E), una nueva droga inhibidora de trombina, con warfarina (C) para la prevención de eventos embólicos sistémicos y accidentes cerebrovasculares (ACV) (*outcome*) en pacientes con fibrilación auricular (paciente, estudio: SPORTIF III). Basándose en estudios previos se fijó una razón de 3,1% de ACV/embolias sistémicas al año con el uso de warfarina en pacientes con fibrilación auricular. El margen de no-inferioridad se fijó en 2%/año con un intervalo de confianza de 95%. El

objetivo primario del estudio es demostrar que el ximelagatran no causa un exceso de 2% o más de ACV o embolias sistémicas al año comparado con warfarina. En base a esto se puede formular las siguientes hipótesis:

$$H_0: \text{Tasa}_E - \text{Tasa}_C > 2\% \text{ al año} \quad H_A: \text{Tasa}_E - \text{Tasa}_C \leq 2\% \text{ al año}$$

Según los resultados del estudio se puede rechazar la hipótesis nula y aceptar la no inferioridad del ximelagatran con respecto a la warfarina en la prevención de ACV-embolias sistémicas. Si no se puede rechazar la hipótesis nula se asume que la hipótesis de inferioridad es razonable.

El método de este tipo de ensayo clínico es más confiable si se realiza una comparación con un control activo adecuadamente probado contra placebo y con un tamaño muestral suficiente para establecer un margen  $\Delta$  adecuado. Sin embargo, este escenario idílico no ocurre con frecuencia, ya que en la realidad clínica lo que frecuentemente se pretende hacer es insertar un nuevo medicamento dentro del mercado como alternativa para la terapia estándar. Además, se ha visto que mientras menos riguroso el diseño del estudio (por ej. margen de no inferioridad muy amplio, tamaño muestral pequeño, poca evidencia de superioridad a placebo del control activo) más probabilidades de que el resultado sea de no inferioridad, y así insertar al mercado un medicamento que puede incluso ser menos efectivo, más riesgoso o más costoso que el control activo.

## PRINCIPALES PROBLEMAS DE LOS ESTUDIOS DE NO INFERIORIDAD

### 1) Sensibilidad del Ensayo (Assay sensitivity):

Este concepto aunque algo difícil de comprender por sí solo tiene gran importancia al analizar la validez de un estudio de no inferioridad. Si nos encontramos con un estudio que concluye la no inferioridad entre dos tratamientos, pero no posee sensibilidad del ensayo dicha conclusión es inválida.

¿Qué determina la sensibilidad del ensayo? Lo principal que hay que evaluar es la evidencia histórica de que el control activo es superior al placebo y lo seguirá siendo en el contexto del estudio de no inferioridad que estamos analizando. Para esto es necesario asegurarse de un adecuado análisis estadístico de estudios realizados en circunstancias similares a las del estudio de no inferioridad.

Lo problemático de este concepto es la limitada cantidad de ensayos clínicos controlados por placebo hechos de forma adecuada, con tamaños

muestrales adecuados, sin (o pocos) sesgos, que sean comparables con otros estudios que busquen el mismo resultado.

## 2) Análisis según Intención de Tratar (ITT) vs Por Protocolo (PP):

Dentro de los estudios de no inferioridad hay dos posibles métodos de análisis de los grupos de estudio, estos son: el análisis ITT y el PP.

El análisis ITT toma en cuenta los resultados del seguimiento completo de todos los pacientes aleatorizados en el estudio, independiente de su cumplimiento con el tratamiento. Este método puede evitar o disminuir las diferencias de tratamiento idealizadas que pueden surgir bajo una premisa poco realista de un 100% de cumplimiento. Sin embargo, surge un dilema al momento de enfrentar este tipo de análisis; la posibilidad de aumentar artificialmente la afirmación de no inferioridad al “diluir” verdaderas diferencias de tratamiento.

El otro tipo de análisis es PP, que se enfoca sólo en pacientes que cumplen el tratamiento o hasta el momento que el cumplimiento cese en cada paciente, en pos de evidenciar diferencias de tratamiento no deseadas. Al igual que en el análisis por ITT, surgen problemas con este método, el grupo de pacientes que cumplen con el tratamiento son un grupo selecto que pueden sesgar favorablemente el resultado ya que muchas veces el no seguimiento puede deberse al simple hecho de la percepción del paciente de que “la terapia no funciona, por lo tanto, dejo de usarla”; estas dificultades aumentan si el cumplimiento difiere entre los distintos grupos de estudio.

Ambos métodos de análisis son aceptados y cada uno tiene sus sesgos; por eso, los nuevos estudios de no inferioridad utilizan ambos tipos de análisis y los comparan: si los resultados son congruentes tanto en el ITT como en el PP se puede concluir con mayor respaldo la no inferioridad.

## 3) Escoger el margen $M$ ( $\Delta$ ):

La importancia de escoger un margen de no inferioridad adecuado (no muy amplio) se da principalmente en esos estudios donde se considera poco ético realizar un estudio controlado por placebo. Lo peor que puede ocurrir es que el margen sea fijado tan amplio que un nuevo tratamiento no mucho mejor que el placebo sea aceptado como no inferior. Es por esto que es mejor inferir de estudios controlados por placebo previos la magnitud de superioridad del control activo sobre el placebo, y el margen  $\Delta$  que se escoja debe ser significativamente menor que este estimado.

## 4) Tamaño muestral:

En los estudios de no inferioridad los tamaños muestrales van estrechamente vinculados con el margen de no inferioridad y el intervalo de confianza que se escoja, ya que si hay verdadera no-inferioridad se debe estudiar un  $N$  suficiente para tener una mayor probabilidad que el intervalo de confianza esté en el lado  $+\Delta$ . Por ejemplo, en la comparación de la eficacia de un nuevo medicamento con el omeprazol para la erradicación del *H. Pylori*, se sabe que la razón de erradicación del omeprazol es de un 85% y para el estudio se determinó un margen  $\Delta$  de 15% (es decir, el medicamento no será inferior al omeprazol si el IC 95% no incluye una inferioridad de 15% relativa al omeprazol). El cálculo del tamaño muestral indica que se requiere de 238 pacientes aleatorizados para realizar un estudio confiable. ¿Por qué usar un margen de 15%? esto parece un poco excesivo, ¿Por qué no usar un 10% o un 5% y asegurarnos de un margen de seguridad más estrecho? Esto es porque para establecer un margen de 10% necesitarías de un  $N = 535$  pacientes, y con  $\Delta$  de 5% necesitaríamos de 2.142 pacientes, lo cual para muchos centros de investigación es un número poco alcanzable. Esta realidad hace que se asuma el riesgo de “dejar pasar” tratamientos que son realmente inferiores y aceptarlos como tratamientos equivalentes.

## 5) *Biocreep*:

El “*biocreep*” es el fenómeno que ocurre cuando un medicamento o tratamiento levemente inferior se convierte en el control activo para la siguiente generación de ensayos de no inferioridad, así hasta que el control activo sea no mucho mejor que el placebo.

## CONCLUSIÓN

En la lectura crítica de un estudio de no inferioridad se debe considerar:

- La superioridad del control activo comparado con placebo es irrefutable, está bien documentada.
- El ensayo de no-inferioridad ha sido conducido en condiciones similares (*assay sensitivity*).
- El margen de no inferioridad debe ser lo suficientemente pequeño para convencernos de que el nuevo tratamiento tiene verdadero valor terapéutico.

En la realidad, muchos estudios de este tipo no cumplen con los criterios necesarios para tener verdadera validez, aumentando la falsa aceptación

de tratamientos inferiores como terapia alternativa. Actualmente, los ensayos clínicos de no inferioridad son fuertemente cuestionados, ya que no necesariamente se basan en el bienestar del paciente, sino en los intereses de la industria farmacéutica,

que busca una posición particular con un tratamiento específico. Se debe tener mucho cuidado en la interpretación de este tipo de trabajos, especialmente en la independencia de los investigadores y los distintos tipos de sesgos ya presentados.

## REFERENCIAS

---

1. Ravina B, Cummings J, McDermott M, Poole RM. Non-inferiority Trials. In Bernard Ravina, et al. (Eds.), *Clinical Trials in Neurology; Design, Conduct, Analysis* New York, NY: Cambridge University Press. (2012). pp 135-145.
2. D'Agostino RB, Massaro JM, & Sullivan LM. Non-inferiority trials: design concepts and issues - the encounters of academic consultants in statistics. *Statistics in Medicine*, 2003; 22: 169-186.
3. Eriksson BI, Dahl OE, Rosencher N, et al. Dabigatran etexilate *versus* enoxaparin for prevention of venous thromboembolism after total hip replacement: a randomised, double-blind, non-inferiority trial. *Lancet* 2007; 370: 949-956.
4. Fleming TR. Current issues in non-inferiority trials. *Statistics in Medicine* 2008; 27: 317-332.
5. Hahn S. Understanding noninferiority trials. *Korean Journal of Pediatrics*, 2012; 55: 403-407.
6. Kaul S, & Diamond GA. *Ann Internal Medicine Academia and Clinic Good Enough: A Primer on the Analysis and Interpretation of Noninferiority Trials*. *Annals of Internal Medicine*, 2006; 145: 62-69.
7. Piaggio G, Elbourne DR., Pocock S, et al. Reporting of noninferiority and equivalence randomized trials: extension of the CONSORT 2010 statement. *JAMA* 2012; 308: 2594-2604.
8. Pocock SJ. The pros and cons of noninferiority trials. *Fundamental & Clinical Pharmacology*, 2003; 17: 483-490.
9. Schumi J, & Wittes JT. Through the looking glass: understanding non-inferiority. *Trials*, 2011; 12: 106.

---

Correspondencia a:  
Dr. David Torres P.  
dtorresp@gmail.com