



CIRUGÍA ESPAÑOLA

www.elsevier.es/cirugia

Carta metodológica

Ensayo clínico de no inferioridad. Características y consideraciones prácticas

Non-inferiority clinical trial. Features and practical considerations

Xavier Serra-Aracil^{a,*} y Marc Fradera^b^a Consorci Corporació Sanitària Parc Taulí CCSPT, Institut d'Investigació i Innovació Parc Taulí I3PT-CERCA, Departamento de Cirugía, Universitat Autònoma de Barcelona, Sabadell, Barcelona, España^b Responsable de la Unitat de Suport a la Recerca (USR), Institut d'Investigació i Innovació Parc Taulí I3PT-CERCA, Sabadell, Barcelona, España**Estudios prospectivos controlados aleatorizados. ¿Cuándo están indicados los estudios de no inferioridad?**

La mayoría de los estudios prospectivos, controlados y aleatorizados (EPCA) son los denominados de superioridad, los cuales tienen como objetivo determinar si una intervención es superior a otra. El hecho de no hallar diferencias estadísticamente significativas entre dos intervenciones no significa que sean similares o equivalentes. Por ello están los ensayos de equivalencia, que tienen como objetivo determinar si una nueva intervención es terapéuticamente similar a otra generalmente existente¹.

Los ensayos de no inferioridad (ENI) evalúan si un nuevo tratamiento no es considerado peor, o inferior, en comparación con el tratamiento habitual. Estos estudios responden a si estamos dispuestos a aceptar una nueva intervención que puede ser clínicamente inferior, hasta un cierto margen, pero que aun así sea beneficiosa para los pacientes y tenga alguna ventaja adicional, como un tratamiento menos agresivo, más seguro y con mejor coste-eficacia².

Los ENI son cada vez más comunes, pero en ocasiones pueden estar mal interpretados. Se necesita una mejor comprensión de las características de un ENI y su interpretación^{3,4}. La declaración original CONSolidated Standards of Reporting Trials (CONSORT) se desarrolló para ayudar a solucionar el problema de la presentación inadecuada de los

EPCA. Para facilitar el desarrollo y la calidad de los ENI se publicó una extensión de la guía CONSORT 2010¹.

La no inferioridad y la equivalencia a veces se utilizan erróneamente de igual manera. Los ensayos de equivalencia se definen como el efecto del tratamiento estando dentro de ambos márgenes de no inferioridad (MNI), por encima y por debajo ($-\Delta$ y $+\Delta$).

La baja calidad de estos estudios puede sesgar sus resultados y producir falsos positivos para la no inferioridad². El objetivo de esta carta es abordar la forma de presentación de los ENI.

Características principales de los ENI

Los ENI deben seguir el checklist propuesto por la guía CONSORT adaptado para este tipo de estudios¹. Existen una serie de características que marcan su identidad y que todos ellos deben reflejar de forma clara para asegurar la calidad del ENI²:

1. *El margen de no inferioridad (MNI)*. Es el valor que permite que un nuevo tratamiento sea considerado «aceptablemente peor».
2. Definir los tipos de análisis de las poblaciones.
3. Definir la consistencia de los intervalos de confianza (IC).
4. *El manejo de datos perdidos (missing)*. Algunos datos perdidos son inevitables, pero en los ENI los datos perdidos pueden sesgar hacia la demostración de la no inferioridad.

* Autor para correspondencia.

Correo electrónico: xserraa@gmail.com (X. Serra-Aracil).

<https://doi.org/10.1016/j.ciresp.2024.07.007>

0009-739X/© 2024 Publicado por Elsevier España, S.L.U. a nombre de AEC.

Cómo informar los resultados de los ENI

Los EPCA de superioridad pueden clasificarse como positivos o negativos según si se rechaza la hipótesis nula (es decir, $p < 0,05$). Por contra, los ENI pueden interpretarse como no inferiores, no concluyentes, inferiores o superiores según la ubicación del IC del 95%⁴.

En la [figura 1](#) se describen todas las posibilidades de los resultados de un ENI en función de dónde se localiza el IC del 95%. En la literatura, a veces se ha utilizado el término no inferioridad cuando la diferencia entre dos tratamientos es no concluyente sin referencia a un margen de no inferioridad. Este uso es inapropiado. Si el ensayo tuvo poca potencia, la diferencia real aún podría ser sustancial.

En los ENI, crear una figura con la posición de los IC del 95% ([fig. 1](#)) con el MNI ayuda a los lectores a interpretar el resultado¹. Así, como se observa en la [figura 1](#), los casos C, E y F fueron estadísticamente significativamente inferiores, pero sin embargo en términos de ENI, C fue no inferior, E no concluyente y F inferior, en relación al MNI⁵.

La no inferioridad puede evaluarse ya sea como una diferencia relativa o absoluta. Ambas formas tienen sus ventajas y desventajas. La diferencia absoluta a menudo es

la más relevante en los resultados de los pacientes, ya que refleja un cambio en la reducción o mejoría asociada con el tratamiento quirúrgico estándar. Puede compararse más fácilmente con otros efectos del tratamiento.

La diferencia relativa puede ser particularmente útil para comprender cómo se compara el nuevo tratamiento con el tratamiento convencional en términos del efecto total del tratamiento con la ausencia de tratamiento³.

Cómo definir el margen de no inferioridad (Δ)

Conocer el MNI, también llamado delta (Δ), es el principal desafío en el diseño de un ENI. Su elección es fundamental para llegar a conclusiones adecuadas. El MNI debería depender de la importancia del resultado final y la magnitud del nuevo tratamiento¹.

A pesar de ser un componente clave en el diseño y la interpretación de los ENI, las recomendaciones para la selección de los MNI no son claras. Una afirmación frecuentemente utilizada en algunos artículos es «la elección del margen cuando fue clínicamente aceptable». Este hecho no tiene suficiente consistencia para la justificación del MNI. En general, el margen se debe elegir sobre una base clínica,

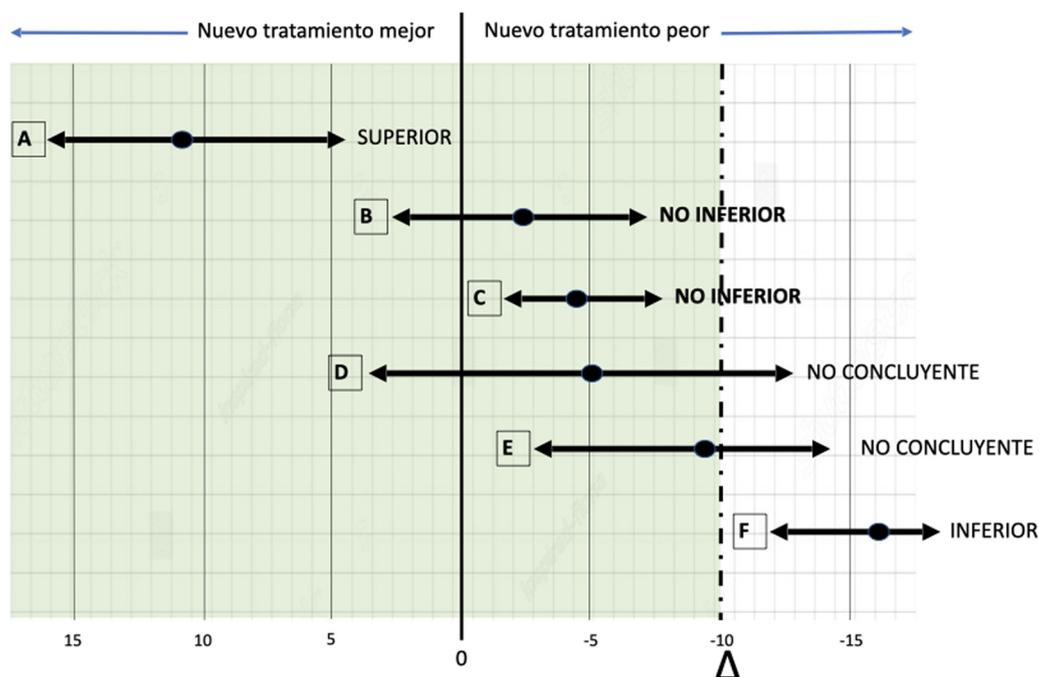


Figura 1 – Localización de los posibles resultados de los intervalos de confianza del 95%. La línea continua es el 0, que indica la significación estadística cuando todo el intervalo se encuentra a un lado u otro del 0. La línea entrecortada es la que indica el margen de no inferioridad ($\Delta = -10$). La zona sombreada hacia la izquierda indica el área de no inferioridad.

A: si el IC se encuentra totalmente a la izquierda de cero, el nuevo tratamiento es superior y por descontado, no inferior. B: si el IC se encuentra a la izquierda de Δ e incluye cero, el nuevo tratamiento no es inferior pero no ha demostrado ser superior. C: si el IC se encuentra completamente a la izquierda de Δ y completamente a la derecha de cero, el nuevo tratamiento no es inferior en el sentido ya definido, pero también es inferior en el sentido de que se excluye una diferencia de tratamiento nula, por lo que será estadísticamente inferior. Esta circunstancia desconcertante es rara, porque requiere un tamaño de muestra muy grande. También puede deberse a un margen de no inferioridad demasiado amplio. D: si el IC incluye Δ y cero, la diferencia no es significativa pero el resultado sobre la no inferioridad no es concluyente. E: si el IC incluye Δ y está completamente a la derecha de cero, la diferencia es estadísticamente significativa pero el resultado no es concluyente con respecto a una posible inferioridad de magnitud Δ o peor. F: si el IC está totalmente a la derecha de Δ , el nuevo tratamiento es inferior¹.

ESTUDIO DE NO INFERIORIDAD CON VARIABLE PRINCIPAL CONTINUA

Cálculo basado en la fórmula:

d: margen de no inferioridad. α : riesgo alfa, tipo I. β : riesgo beta, tipo II

$$n = f(\alpha, \beta) \times 2 \times \sigma / d$$

donde σ es la desviación estándar, y

$$f(\alpha, \beta) = [\Phi(\alpha) + \Phi(\beta)]$$

Φ es la función de distribución acumulativa de una desviación normal estandarizada.

ESTUDIO DE NO INFERIORIDAD CON VARIABLE PRINCIPAL DICOTÓMICA

Cálculo basado en la fórmula:

d: margen de no inferioridad. α : riesgo alfa, tipo I. β : riesgo beta, tipo II

$$n = f(\alpha, \beta) \times [\pi_1 \times (100 - \pi_1) + \pi_2 \times (100 - \pi_2)] / (\pi_1 - \pi_2 - d)$$

donde π_1 y π_2 son el verdadero porcentaje de "éxito" en el grupo de tratamiento estándar y experimental respectivamente, y

$$f(\alpha, \beta) = [\Phi(\alpha) + \Phi(\beta)]$$

Φ es la función de distribución acumulativa de una desviación normal estandarizada.

Figura 2 – Fórmulas del cálculo de la muestra de estudios de no inferioridad.

incorporando evidencia estadística de metaanálisis u otros estudios clínicos de calidad⁶. Si la elección del margen se basa en un grupo de expertos clínicos, se debe proporcionar información sobre cuántos expertos estuvieron involucrados y cuántos consideraron aceptable la elección del margen (como en los procesos Delphi)².

Cómo realizar el cálculo muestral

Difiere considerablemente respecto a los estudios de superioridad, ya que están directamente definidos por el MNI (Δ). Si Δ es demasiado grande, habrá un riesgo muy alto de aceptar tanto un tratamiento verdaderamente inferior como no inferior. Esta situación es especialmente relevante para resultados importantes, como la mortalidad. Por otro lado, definir un Δ muy pequeño puede producir resultados inconclusos, requiriendo un ensayo extremadamente grande.

Es necesario conocer el valor de la potencia estadística ($1 - \beta$), que se asigna al estudio. La mayoría de los investigadores asignan un valor $\beta = 0,2$, por lo que la potencia es de 0,8. Esta puede ser mayor si se asume que el nuevo tratamiento es más efectivo que el tratamiento de referencia, o menor si se asume que es menos efectivo. También el riesgo tipo I, alfa (α), que debe estar relacionado según lo especificado en la hipótesis de no inferioridad. El tipo de contraste bilateral proporciona información adicional, en particular para la situación en la que el nuevo tratamiento sea superior al tratamiento habitual.

En la figura 2 se muestran las fórmulas de cálculo del tamaño de la muestra para estudios de no inferioridad para una variable principal con dos categorías (2A) o una variable continua (2B). En los siguientes enlaces se puede acceder a las calculadoras online para ENI con ambos tipos de variables: <<https://www.sealedenvelope.com/power/continuous-noninferior/>>⁷ y <<https://www.sealedenvelope.com/power/binary-noninferior/>>⁸.

Especificar el tipo de análisis de la población que está siendo evaluada

Se debe considerar que tanto el análisis por intención de tratar (ITT, por sus siglas en inglés) como el de por protocolo

cumplan la no inferioridad. Por lo general, la población del análisis ITT es la más apropiada, pero como se mencionó anteriormente, esto puede subestimar la diferencia real entre tratamientos, y es importante describir también el cumplimiento con el tratamiento en ambos grupos³.

¿Debe realizarse un análisis intermedio? Detención del estudio

En ensayos de superioridad, cuando un análisis intermedio muestra evidencia clara de la eficacia del nuevo tratamiento, puede considerarse éticamente incorrecto continuar el ensayo y negar el nuevo tratamiento efectivo al grupo control. En cambio, en ENI, si se demuestra la no inferioridad para el resultado principal antes de la finalización del ensayo, hay menos necesidad ética de detener el ensayo porque el grupo de control ya está recibiendo el tratamiento habitual. Por tal motivo, se sugiere finalizar los ENI y poder definir concretamente dónde se sitúa el IC.

Comentario final

Se deben describir claramente las características principales de los ENI que hemos descrito. La ausencia en su definición puede llevar a conclusiones falsas², especialmente en aquellos ENI con resultados no concluyentes, que pueden malinterpretar los verdaderos resultados del ensayo.

BIBLIOGRAFÍA

- Piaggio G, Elbourne DR, Pocock SJ, Evans SJ, Altman DG, CONSORT Group. Reporting of noninferiority and equivalence randomized trials: Extension of the CONSORT 2010 statement. JAMA. 2012;308:2594–604. <http://dx.doi.org/10.1001/jama.2012.87802>.
- Rehal S, Morris TP, Fielding K, Carpenter JR, Phillips PP. Non-inferiority trials: Are they inferior? A systematic review of reporting in major medical journals. BMJ Open. 2016;6:e012594. <http://dx.doi.org/10.1136/bmjopen-2016-012594>.

3. Cuzick J, Sasieni P. Interpreting the results of noninferiority trials — a review. *Br J Cancer*. 2022;127:1755–9. <http://dx.doi.org/10.1038/s41416-022-01937-w>.
4. Parsyan A, Marini W, Fazelzad R, Moher D, McCreedy D. Current issues in conduct and reporting of noninferiority randomized controlled trials in surgical management of cancer patients. *Ann Surg Oncol*. 2021;28:39–47. <http://dx.doi.org/10.1245/s10434-020-08575-7>.
5. Acuna SA, Dossa F, Baxter NN. Frequency of misinterpretation of inconclusive noninferiority trials: The case of the laparoscopic vs open resection for rectal cancer trials. *JAMA Surg*. 2019;154:90–2. <http://dx.doi.org/10.1001/jamasurg.2018.3222>.
6. Acuna SA, Chesney TR, Amarasekera ST, Baxter NN. Defining non-inferiority margins for quality of surgical resection for rectal cancer: A Delphi consensus study. *Ann Surg Oncol*. 2018;25:3171–8. <http://dx.doi.org/10.1245/s10434-018-6639-7>.
7. Sealed Envelope Ltd. 2012. Power calculator for continuous outcome non-inferiority trial [consultado 5 May 2024]. Disponible en: <https://www.sealedenvelope.com/power/continuous-noninferior/>
8. Sealed Envelope Ltd. 2012. Power calculator for binary outcome non-inferiority trial [consultado 6 May 2024]. Disponible en: <https://www.sealedenvelope.com/power/binary-noninferior/>