

## **(Resumen de la presentación realizada en la reunión de la Comisión de Bioética de la Fundación Femeba-Buenos Aires 10 de Octubre 2013)**

La medicina basada en evidencia, nos exige antes de utilizar un medicamento, que el mismo haya demostrado una adecuada relación beneficio /riesgo, y que los ensayos clínicos se presenten en base a resultados con variables clínicamente relevantes o duras (muertes, infarto de miocardio, fracturas etc.) resultando de menor importancia aquellas con variables subrogadas (glucemia, presión arterial, hemoglobina glicosilada etc.). La fármaco-economía ha venido a profundizar aún más éstos temas con la necesidad de completar estudios de costo-efectividad<sup>1</sup>. La era de oro de la farmacología quedó atrás y en la actualidad se hace más difícil obtener medicamentos innovadores, esto trae aparejado una tensión entre resultados científicos y las necesidades económicas, en donde la industria farmacéutica necesita desarrollar sus productos en circunstancias no tan favorables y de alguna manera igualmente, colocar sus productos en el mercado para su venta.

Los estudios clínicos aleatorizados que tienen como objetivo detectar diferencias en los tratamientos que un nuevo medicamento en comparación con un placebo u otro medicamento activo, se denominan de “**Superioridad**”<sup>2</sup> Existen además dos clases de ensayos clínicos denominados de “**Equivalencia**” y de “**No inferioridad**” (ver tabla 1), que pasamos a describir: - Los ensayos de **equivalencia** demuestran que no existen diferencias significativas clínicamente relevantes, entre los tratamientos, no son superiores ni inferiores, tampoco son iguales, pero se encuentran dentro de un margen de superioridad y de inferioridad pre-establecidos. Fueron muy usados en la década del 90, en estudios de comparación de eficacia de los llamados homólogos terapéuticos (me too), que permiten realizar el reemplazo de un medicamento por otro, para una misma indicación. Para estos estudios es necesario que los fármacos cumplan una serie de requisitos: poseer las mismas indicaciones aprobadas, igual mecanismo de acción o pertenecer al mismo grupo farmacológico y haber demostrado eficacia y seguridad entre otras razones<sup>3</sup>. No debemos confundirnos con los análisis de **biodisponibilidad** donde se compara un medicamento original versus un genérico para demostrar algunas propiedades

---

<sup>1</sup> Cornel Pater Equivalence and noninferiority trials – are they viable alternatives for registration of new drugs? (III) *Current . Trials in Cardiovascular Medicine* 2004, 5:8 doi:10.1186/1468-6708-5-8.

<sup>2</sup> *Ensayos clínicos de superioridad, equivalencia y no inferioridad* [www.revistaacofar.com](http://www.revistaacofar.com) › Inicio › I + D + I 26/07/2013.

<sup>3</sup> Gonzalez Soto MI; Casamayor Lazaro B, Garcia del Valle R et al. Equivalencia Terapéutica: Importancia en la Práctica Clínica. *Boletín de Farmacoterapéutica de Castilla-La Mancha* 2012; Vol XIII(4):1-8.

farmacocinéticas (relaciones de áreas bajo la curva, porcentajes y tiempos de absorción, concentración sanguínea), y que nos permitirá o no su intercambiabilidad terapéutica<sup>4</sup>.

Por otra parte, los ensayos clínicos de **no inferioridad** tienden a demostrar, que un tratamiento no es significativamente peor a otro. En realidad solamente se pretende demostrar que no sea peor, y se propone cual es la diferencia que podemos aceptar entre ambos tratamientos para que dicha diferencia sea considerada no relevante clínicamente<sup>5</sup>.

---

<sup>4</sup> Olga Delgado Sánchez <sup>a</sup>, Francesc Puigventós Latorre <sup>a</sup>, Manel Pinteño Blanco <sup>b</sup>, Pere Ventayol Bosch Equivalencia terapéutica: concepto y niveles de evidencia Vol. 129. Núm. 19. 24 Noviembre 2007.

<sup>5</sup> Estudios de no inferioridad ¿Cuánta evidencia? Elard Walter Quispe Mena Salud y Fármacos.10 de Agosto de 2010.

**Tabla 1**

<p style="text-align: center;"><b>Definición de equivalencia y no inferioridad</b></p> <ul style="list-style-type: none"><li>• Ensayos de equivalencia: demuestran que no existen diferencias significativas entre los tratamientos</li><li>• Ensayos de no inferioridad: demuestran que un tratamiento no es significativamente peor a otro tratamiento</li></ul>
--

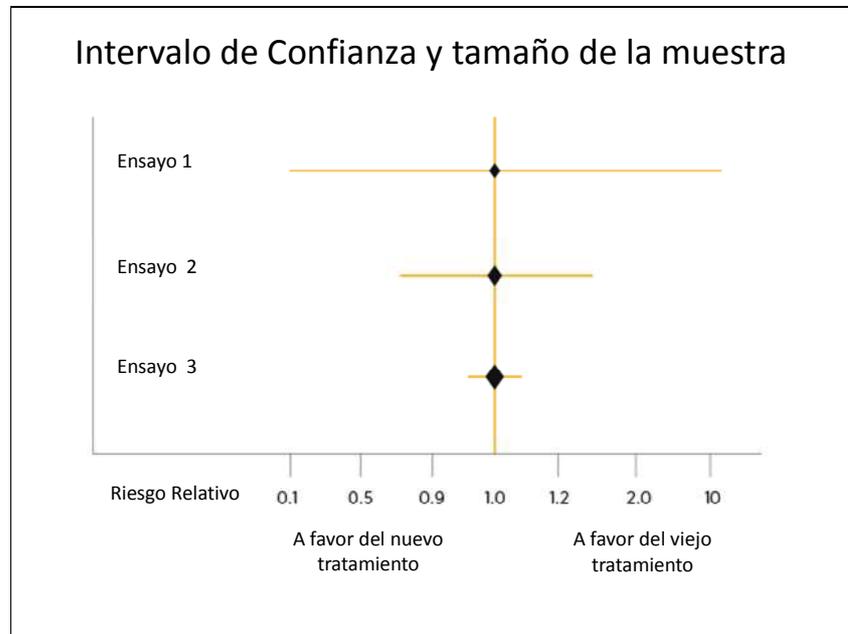
## **CONSIDERACIONES METODOLÓGICAS GENERALES**

**Número de la muestra e intervalo de confianza.** El número de pacientes de un ensayo depende, de lo que nosotros esperamos obtener de un determinado evento, y que éste resulte estadísticamente significativo. Si suponemos que el medicamento puede bajar un 40 % los eventos a estudiar (muy efectivo), necesitaremos mucho menos pacientes, que si solo esperamos un efecto del 7%. El intervalo de confianza es el margen de seguridad de un determinado valor, que nos asegura que en un 95 % (90 % o 99 %) de los casos existe la probabilidad que un dato experimental sea verdadero y no dependiente del azar. A mayor número de nuestra más estrecho los límites del índice de confianza, y más confiables los datos obtenidos<sup>6</sup> (Ver gráfico 1).

---

<sup>6</sup> Carlos Tajer y Hernán Doval COMO LEER UN ENSAYO CLINICO PARA INCORPORARLO A NUESTRO CONOCIMIENTO MEDICO Y APLICARLO EN LA PRACTICA [www.gedic.com.ar/publicaciones/ Como\\_Leer\\_un\\_Ensayo\\_Clinico.doc](http://www.gedic.com.ar/publicaciones/Como_Leer_un_Ensayo_Clinico.doc).

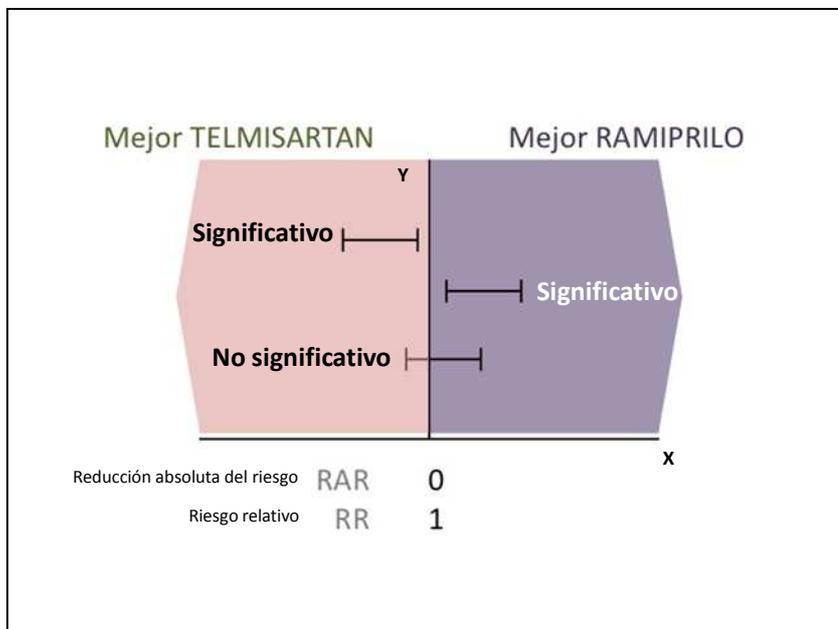
**Gráfico 1**



La representación gráfica, de los resultados en los ensayos clínicos se realiza sobre coordenadas cartesianas, en el eje de la X (abscisas) se representan las diferencias de eventos estudiados como Reducción del Riesgo Absoluto (RRA), o en términos de Riesgo Relativo (RR). La Reducción del Riesgo Relativo ( $RRR = 1 - RR$ ) representa la disminución proporcional de eventos o acontecimientos entre el grupo experimental y el control, expresados en porcentaje (%). Es también importante considerar el NNT (Número Necesario de Tratar) representado por el valor de  $1/RRA$  y cuyo significado representa el número de pacientes a tratar para obtener un beneficio en la variable clínica estudiada.

Cuando el intervalo de confianza (IC 95 %) sobrepasa los valores de 0 para las diferencias absolutas, o el valor de 1 en los riesgos relativos, consideramos que las diferencias no son significativas, ya que existe la probabilidad de que un valor experimental resulte igual a cero en RAR o a uno en valores de RR (Ver gráfico 2).

Gráfico 2



### DETERMINACIÓN DEL MARGEN DE “NO INFERIORIDAD”

El margen de no inferioridad puede establecerse por dos criterios: clínico, y por métodos estadísticos.

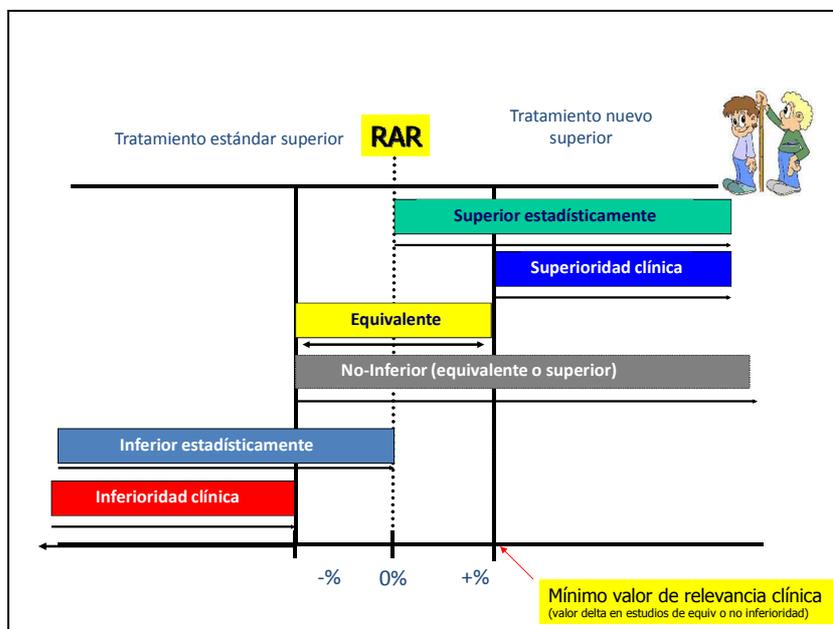
**El método clínico** establece un margen, donde las diferencias entre el efecto del nuevo medicamento versus el control activo, no resultan clínicamente relevantes, por ejemplo: una diferencia de 5 o 10 mm de Hg de presión sanguínea, o 10 mm de una escala analógica de dolor. Este control activo es el que demostró ser superior en ensayos previos, en meta análisis o revisiones sistemáticas frente al placebo.

**El método estadístico**, se sustenta sobre el análisis estadístico de diferencias existentes entre el tratamiento estándar y el placebo. En el nuevo ensayo de “no inferioridad”, el estudio se realiza comparando el medicamento a investigar versus un control activo, considerando el comportamiento histórico del control activo de los ensayos previos frente a placebo.

Teniendo en cuenta las diferencias entre los ensayos clínicos de superioridad, equivalencia y no inferioridad, y los márgenes (valores delta positivos y negativos) propuestos, podemos

observar distintas posibilidades comparativas resultantes entre un tratamiento nuevo versus su comparador (Ver gráfico 3).

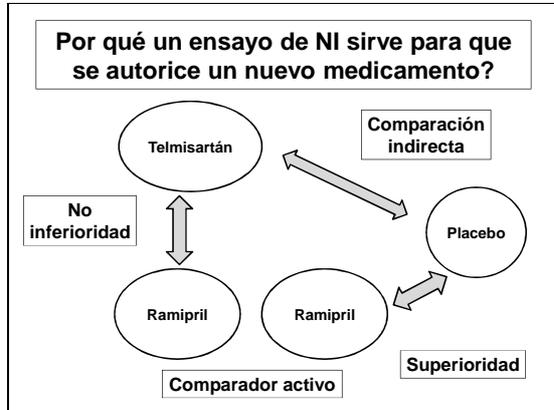
**Gráfico 3**



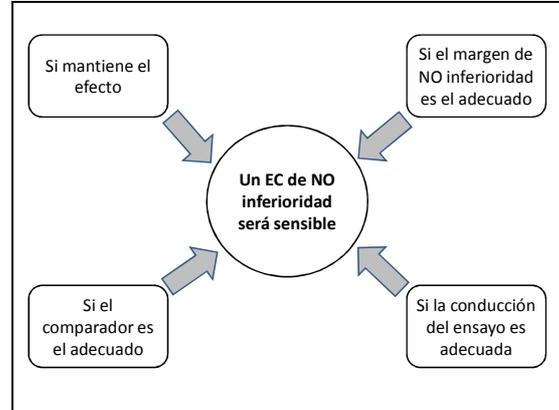
Los ensayos clínicos de no inferioridad, no poseen la misma sensibilidad que tienen los ensayos de superioridad. Un estudio posee sensibilidad, cuando el mismo es capaz de distinguir diferencias de eficacia, como ocurre en los ensayos de superioridad; en cambio en los estudios de no inferioridad, en el que dos medicamentos resulten equivalentes, esto no implica que ambos sean eficaces, incluso pueden ser ineficaces. Además la sensibilidad en los ensayos de no inferioridad se valora en forma indirecta, por la evidencia histórica del control activo versus placebo, el correcto diseño y el seguimiento del estudio. Sus conclusiones están soportadas sobre cuatro pilares: (Ver gráfico 4 y 5)

- 1- Si el comparador(control activo) es el adecuado,
- 2- Si mantiene el mismo efecto, o sea la misma actividad que la observada en los experimentos contra placebo, manteniendo la similitud del diseño
- 3-Si el margen de no inferioridad es el adecuado, y 4- Si la conducción del ensayo fue adecuada

**Grafico 4**



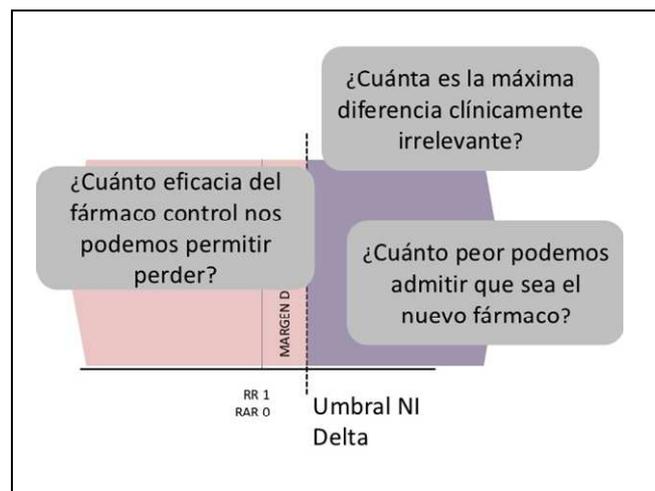
**Gráfico 5**



En los ensayos de no inferioridad, para que un medicamento sea autorizado, es suficiente con que demuestre, que no es inferior al control activo, interpretando en forma indirecta, que el control activo tuvo el mismo comportamiento histórico con el placebo; desde este punto de vista estos ensayos poseen una importante debilidad metodológica<sup>1-5</sup>.

Por lo tanto antes de colocar el margen de no inferioridad (umbral delta) nos debemos preguntar 1- ¿Cuánta eficacia del fármaco control nos podemos permitir perder? 2- ¿Cuánto peor podemos admitir que sea el nuevo fármaco? Y 3- ¿Cuánta es la máxima diferencia clínicamente irrelevante? (Ver gráfico 6).

**Grafico 6**



¿Qué criterio se debe adoptar al seleccionar el margen terapéutico límite para establecer la no inferioridad?, a fin de aceptar que el nuevo tratamiento no sea inferior al tratamiento

previo? Se han establecido varios criterios: 1- Nunca el margen de no inferioridad puede superar el límite superior del intervalo de confianza del efecto de los ensayos previos, considerando que el nuevo tratamiento también es superior al placebo. 2-Considerar que la nueva intervención conserve como límite mínimo el 50 o el 80 % del efecto de la clásica.

Sin embargo en muchos trabajos observamos que dicho margen se ubica en un 50% (en RR) en relación al comparador, cuando debería ser considerado razonable que el valor del margen de no inferioridad se ubique entre el 5 al 20 % de la eficacia de referencia.

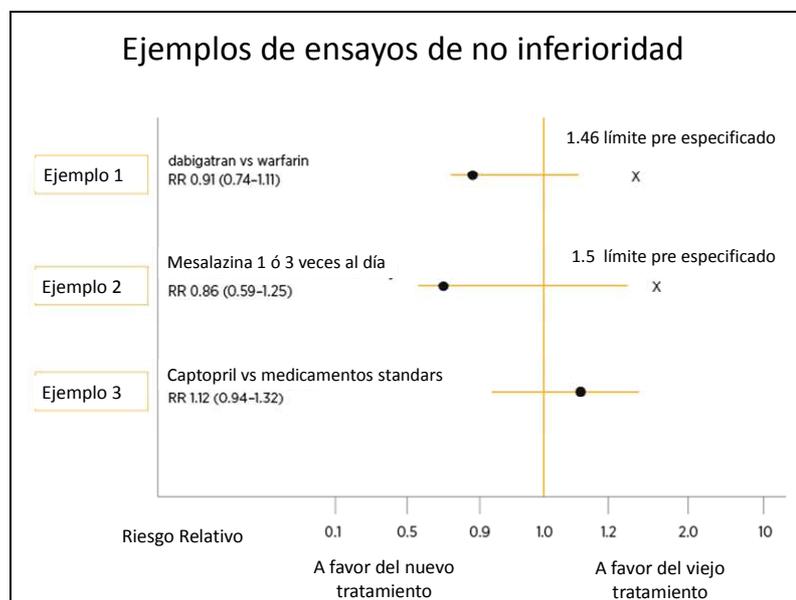
En el estudio de Dabigatrán vs warfarina (estudio RE-LY)<sup>7</sup>, el margen de no inferioridad fue colocado en un riesgo relativo de 1,46, se permitió al dabigatrán otorgarle un 46 % más de eventos adversos (accidentes cerebrovasculares y embolismo sistémico), y considerarlo clínicamente irrelevante, aunque consideramos que es inaceptable. Así cuando se utilizó dabigatrán en dosis de 110 mg cada doce horas el dato observado versus warfarina en riesgo relativo, RR: 0.91 (IC 95 % 0.74-1.11) fue positivo para no inferioridad, con el extremo superior aceptable del 11 % para las variables estudiadas. El borde superior 1.11 no superaba el margen de no inferioridad establecido de 1.46.

En el segundo ejemplo donde se estudiaron los resultados del tratamiento con mesalazina en la remisión de colitis ulcerosa, después del primer episodio, se observó que una sola dosis de mesalazina no era inferior a tres dosis diarias. El margen o valor delta de no inferioridad fue establecido en 1.5 del riesgo relativo; el estudio arrojó un RR:0.86 (IC 95 % 0.59-1.25), valor positivo para no inferioridad. El tercer ejemplo donde se estudió la eficacia del captopril versus otros antihipertensivos en los desenlaces combinados de ACV, IM y muerte cardiovascular, el riesgo relativo fue de 1.12 (IC 95% 0.94-1.32); en este caso se reclamó equivalencia, pero fue inaceptable considerar un 32 % de aumento de los desenlaces como equivalentes y el no haber presentado con anterioridad el margen superior de equivalencia (Ver gráfico 7).

---

<sup>7</sup> Stuart J. Connolly, M.D., Michael D. Ezekowitz, M.B., Ch.B., D.Phil., Salim Yusuf, et al Dabigatran versus Warfarin in Patients with Atrial Fibrillation N Engl J Med 2009;361:1139-51.

## Gráfico 7



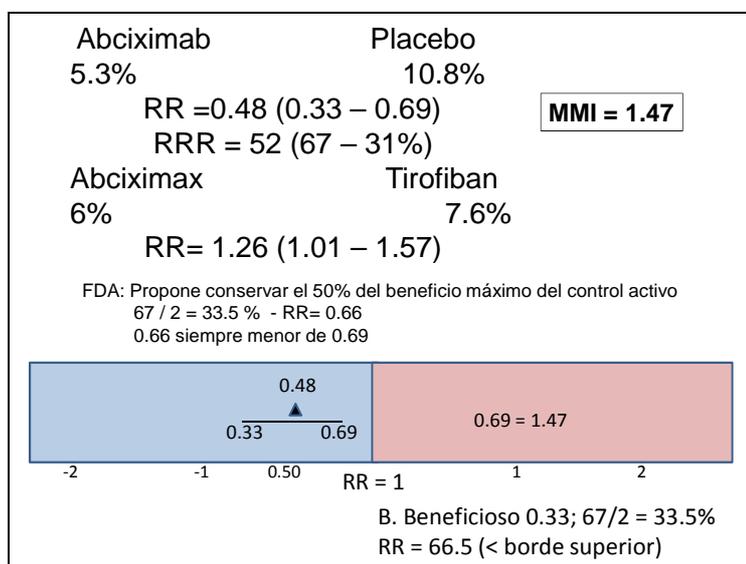
### ESTIMACIÓN DEL MARGEN DE “NO INFERIORIDAD”

Con el objetivo de comprender mejor los criterios estadísticos y o clínicos de cómo se debe establecer el margen de no inferioridad (valor delta de no inferioridad), analicemos el siguiente estudio<sup>8</sup>, donde se comparó la efectividad del tirofiban versus abciximab, en angioplastia programada con stent, y análisis de variables combinadas de muerte, infarto de miocardio y necesidad de revascularización. En estudios previos el abciximab (control activo) fue superior a placebo ,5.3% eventos versus 10.8 % para el placebo con hazard ratio, HR= 0.48 (IC 95 % 0.33-0.69). Se aceptó un RR superior para tirofiban en relación a abciximab de 1.47 (margen de no inferioridad) conservando el 50 % del efecto anterior de abciximab respecto al placebo. El resultado de no inferioridad fue desfavorable para tirofiban, ya que se obtuvieron 7.6 % de eventos para tirofiban y 6 % de eventos para

<sup>8</sup> CD Tajer Estudios de equivalencia y no inferioridad – gedic [www.gedic.com.ar/publicaciones/Equivalencia\\_y\\_no\\_inferioridad.doc](http://www.gedic.com.ar/publicaciones/Equivalencia_y_no_inferioridad.doc).

abciximab con un HR=1.57, que cruzaba el margen de no inferioridad establecido. No existen cálculo matemático o formulas para establecer el valor delta de no inferioridad, quedando a criterio clínico ésta decisión. La FDA propone que el borde inferior(efecto beneficioso máximo) en este estudio del 0.33 %,equivalente a un RRR del 67 %, a fin de conservar la mitad de éste beneficio, la reducción de riesgo no debe ser menor a  $67/2=33.5\%$  o sea un  $RR=0.66$ .Este nuevo margen debe siempre ser inferior al margen superior observado(en éste caso de 0.69).Por lo tanto si aceptamos 0.66 y ahora lo relacionamos al efecto promedio central de 0.48; $0.66/0.48$ , da aproximadamente  $1.4^8$  (Ver gráfico 8).

### Grafico 8

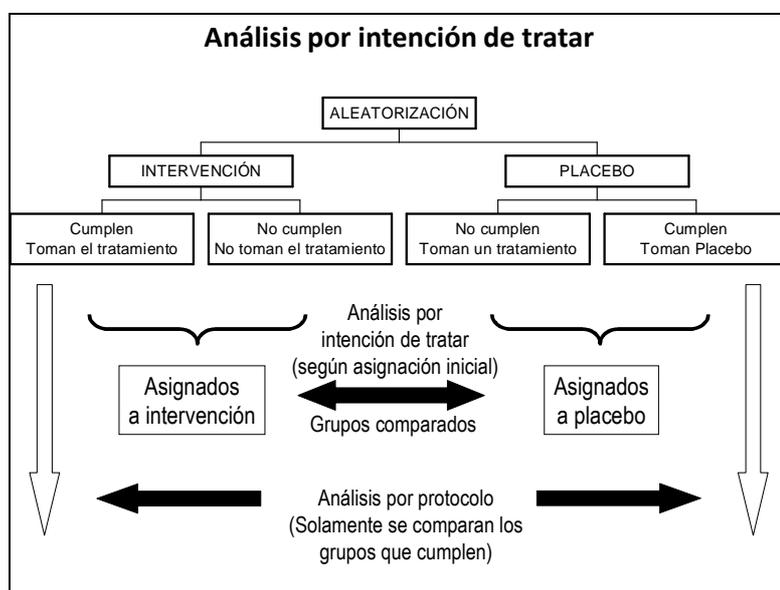


La magnitud del margen de no inferioridad es extremadamente sensible. Cuanto más amplio sea el delta, menor número de muestra se necesitará para demostrar no inferioridad; si el margen es del 13 % se necesitará 25.000 pacientes, si el margen resulta del 8% el número de pacientes requeridos será de aproximadamente de 52.000 .A mayor margen, mayores posibilidades de error en demostrar no inferioridad cuando en realidad hay inferioridad.

**Seguimiento de los ensayos** El análisis de los resultados por intención de tratar (ITT) es ampliamente reconocido como el análisis más importante en los ensayos de superioridad, por que el procedimiento es adherirse a la aleatorización, método conservador. Si a pesar de abandonos o exclusiones (aceptado no más del 20%) después de la aleatorización, los resultados demuestran diferencias significativas, el ensayo de superioridad es más

convinciente. Desafortunadamente el análisis por intención de tratar en los ensayos de no inferioridad, tienden a sesgos con tendencia a la equivalencia, en donde un resultado de inferioridad aparezca como un resultado falso de no inferioridad. Los análisis por protocolo, donde solamente se consideran los pacientes que cumplen con los tratamientos y no se consideran los abandonos o exclusiones, pueden en realidad demostrar la real equivalencia de no inferioridad, pero siempre teniendo en cuenta, qué variables se están estudiando, así por ejemplo, si el estudio es un análisis de sobrevida, y suprimimos pacientes por efecto medicamentoso, por producir insuficiencia cardíaca, estamos cometiendo un sesgo muy importante, ya que la insuficiencia cardíaca es un factor de riesgo de mortalidad. De allí la recomendación siempre de realizar ambos procedimientos en el seguimiento en los ensayos de no inferioridad.<sup>9, 10.</sup>

**Gráfico 9**



## CONCLUSIONES

Existen, como queda demostrado problemas inherentes a los ensayos clínicos de “no inferioridad”, que los hacen menos creíbles que aquellos de superioridad, o los realizados

<sup>9</sup> Steven M Snapinn Noninferiority trials Curr Control Trials Cardiovasc Med. 2000; 1(1): 19–21. Published online 2000 July 31. doi: 10.1186/cvm-1-1-019.

<sup>10</sup> Ben Ewald Making sense of equivalence and non-inferiority trials Australian Prescriber 170 VOLUME 36 : NUMBER 5 : OCTOBER 2013.

contra placebo. La argumentación para su realización está sustentada débilmente, en que no es ético el uso de placebo, lo cual es cierto. Aunque se acepte que estos ensayos no demuestran superioridad, es posible demostrar algunas ventajas como: mejoramiento de la adherencia al tratamiento, el nuevo medicamento resulta con mejor perfil de costo-efectividad etc., etc. Tampoco estos argumentos son claros en la realidad. Además es importante remarcar que en muchos trabajos publicados, no se establece en forma prioritaria y explícita la determinación del margen de no inferioridad, o se establecen márgenes de no inferioridad amplios aceptando porcentajes de mortalidad mayores para el nuevo tratamiento, totalmente inaceptables éticamente.<sup>11</sup>

## **DUDAS METODOLÓGICAS Y ETICAS**

Numerosos autores se han preocupado en analizar aspectos éticos de los ensayos clínicos de equivalencia y no inferioridad. Consideraremos algunos aspectos metodológicos. Especialmente relacionados con el concepto de incertidumbre, principio científico pero también ético, sin el cual no es correcto que se plantee un estudio comparativo (Djulbagovic B, Clarke M. Scientific and ethical issues in equivalence trials. JAMA 2001;285:1206-1208), (Delgado Sanchez O, Puigventós Latorre F, Pinteño Blanco M et al. Equivalencia Terapéutica: Concepto y Niveles de Evidencia. Med Clin.(Barc) 2007;129(19):736-745).

Los principios básicos aceptables para la realización de un ensayo clínico son :1-La intervención debe estar plenamente justificada y 2-El control de comparación debe ser el más adecuado. Si estamos seguros que un tratamiento es superior a otro, no se justifica ningún tipo de ensayo. En cambio , si existe la incertidumbre, acerca de cuál es el mejor tratamiento, se justifica ética y científicamente la realización de un ensayo clínico de superioridad.

---

<sup>11</sup> Garattini S, Bertele V. Non-Inferiority Trials are unethical because they disregard patients' interests. Lancet 2007;370:1875-1877.

En los ensayos clínicos de superioridad, se parte de la hipótesis nula ( $H_0$ ) de la no existencia de diferencias de los dos tratamientos; en cambio en los ensayos de no inferioridad, la hipótesis nula inicial ( $H_0$ ) plantea que un tratamiento es superior a otro, por lo tanto no está justificado el principio de incertidumbre. Por otra parte este aspecto debería quedar claramente expresado en el consentimiento informado. Esta información ,si se presentare siempre en forma explícita ,permitiría al paciente manifestar su autonomía .

En los ensayos clínicos de no inferioridad, se hace difícil justificar algunos aspectos reñidos con la ética, como establecer el margen de no inferioridad en cuanto a variables de morbilidad o mortalidad, ya que no es posible aceptar más un número mayor de muertes o mas efectos adversos serios, aunque se obtengan beneficios secundarios.

Considerando el margen de no inferioridad, se puede llegar a requerir mayor tamaño muestral que en los ensayos de superioridad, lo que supone exponer a mas pacientes, para obtener pobres o nulos beneficios.

Cada vez más frecuentemente se presentan ensayos clínicos denominados de bioarrastre (fenómeno "Biocreep". (Moran J, Solomon P .The interpretation of lack of evidence of a difference in efficacy equivalence trials and the treatment of fungal infections.Critical Care and Resuscitation 2003;5:216-223), donde se aceptan en forma progresiva tratamientos cada vez menos eficaces, donde B es no inferior a A (control activo) y entonces por arrastre C es no inferior a B y así sucesivamente.

Además, los ensayos clínicos de no inferioridad no aportan mejoría clínica, al no demostrar mayor eficacia que el comparador y poseer en forma indirecta capacidad discriminatoria (sensibilidad)y por lo tanto librada a la metodología y responsabilidad del investigador.