

## 4. *Special designs*

### 4.3.- *Bio-equivalence*

# Medical Statistics

José Antonio González y Erik Cobo

Abril 2015

## Med. Stat: Bio-equivalencia



**Hipótesis de interés** que usualmente interesa establecer:

- 'diferencias':**
- El riesgo de bronquitis es mayor en fumadores
  - El nuevo tratamiento es mejor que el clásico
- 'relaciones':**
- Hay correlación positiva entre peso y talla.
  - Hay correlación dosis-efecto.

En estas situaciones, la prueba de significación funciona, pues lo que se desea establecer está en la hipótesis alternativa.

Pero puede interesar establecer '*equivalencia*':

Un tratamiento es 'equivalente' a otro si la diferencia de sus efectos no alcanza un cierto valor que hace relevantes las consecuencias.

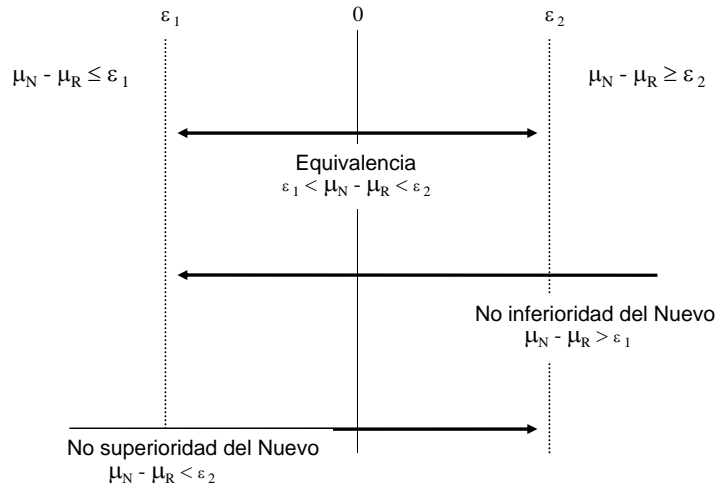
Por ejemplo,  $|\mu_1 - \mu_2| < \varepsilon$

O bien:  $\varepsilon_1 < \mu_1 - \mu_2 < \varepsilon_2$

Nótese:

- 1.- La prueba clásica de diferencias (que tiene en  $H_0$  la igualdad) ya no sirve.
- 2.- El concepto de equivalencia es más amplio que el de igualdad.

Representación gráfica de  $\mu_N - \mu_R$



Margen  $\varepsilon$  de no-inferioridad

Máxima diferencia entre tratamiento experimental y control que se considera clínicamente aceptable o irrelevante y que el ensayo intentará descartar.

Debe asegurar que el efecto del nuevo tratamiento es distinto del de placebo y que mantiene una parte del efecto del control.

No puede ser mayor que la mínima diferencia esperable entre el control activo y placebo, en las mismas condiciones del ensayo (p.ej., un tercio, la mitad ...).

Sólo puede estimarse de forma fiable a partir de EC previos de diseño similar que tuvieron resultados consistentes y demostrada sensibilidad

## Inadecuación de la prueba 'clásica'.

La prueba clásica la usaríamos de la siguiente forma:

$$\begin{cases} H_0: \mu_N = \mu_R & \text{o bien} & \begin{cases} H_0: \mu_N - \mu_R = 0 \\ H_A: \mu_N - \mu_R \neq 0 \end{cases} \end{cases} \rightarrow \begin{cases} \text{Los efectos son iguales, idénticos.} \\ \text{Los fármc. no son intercambiables} \end{cases}$$

Se pone a prueba, por ejemplo, mediante el estadístico  $\bar{t} = (\bar{X}_1 - \bar{X}_2) / S\sqrt{(2/n)}$

Si  $|\bar{t}| < |t_{(\alpha/2, v)}| \rightarrow$  Nada se opone a aceptar  $H_0 \rightarrow$  Los fármacos son intercambiables

Si  $|\bar{t}| > |t_{(\alpha/2, v)}| \rightarrow$  Se rechaza  $H_0 \rightarrow$  Los fármc. no son intercambiables

Nótese las situaciones 'perversas' siguientes:

- 1.- Se desconoce el riesgo  $\beta$ .
- 2.- A medida que aumenta la información que proporciona el estudio, disminuye el valor de  $S\sqrt{(2/n)}$  y aumentan las ocasiones en las que se rechazará la hipótesis nula de interés.

La primera dificultad podría ser solventada con un cálculo de potencia, pero la segunda provoca que estudios 'mejores' conduzcan mas veces a 'no equivalencia'

Prueba doblemente unilateral (PDU) o *Two-one-sided test*

$$\begin{cases} H_0: \mu_N - \mu_R \leq \varepsilon_1 \quad \text{ó} \quad \mu_N - \mu_R \geq \varepsilon_2 \\ H_A: \varepsilon_1 < \mu_N - \mu_R < \varepsilon_2 \end{cases} \rightarrow \begin{cases} \text{Diferencias relevantes} \\ \text{Igualdad + diferencias irrelevantes} \end{cases} \rightarrow \begin{cases} \text{No equivalencia} \\ \text{Equivalencia} \end{cases}$$

donde  $\varepsilon_1$  y  $\varepsilon_2$  delimitan el *intervalo de equivalencia* con valores irrelevantes.  
Usualmente  $\varepsilon_1 = -\varepsilon_2$ .

Este contraste se puede plantear en forma de dos contrastes unilaterales simultáneos:

$$\begin{cases} H_{01}: \mu_N - \mu_R \leq \varepsilon_1 \\ H_{A1}: \mu_N - \mu_R > \varepsilon_1 \end{cases} \quad \text{y} \quad \begin{cases} H_{02}: \mu_N - \mu_R \geq \varepsilon_2 \\ H_{A2}: \mu_N - \mu_R < \varepsilon_2 \end{cases} \rightarrow \begin{cases} \text{No equivalencia} \\ \text{Equivalencia (si } H_{A1} \cap H_{A2}) \end{cases}$$

que se resuelven mediante:

$$t_1 = [(\bar{x}_1 - \bar{x}_2) - \varepsilon_1] / s\sqrt{(2/n)} \quad \text{y} \quad t_2 = [(\bar{x}_1 - \bar{x}_2) - \varepsilon_2] / s\sqrt{(2/n)}$$

Si  $t_1 > t_{(1-\alpha, v)} \rightarrow \mu_N - \mu_R > \varepsilon_1$  y si  $t_2 < t_{(\alpha, v)} \rightarrow \mu_N - \mu_R < \varepsilon_2$

y si ambas pruebas unilaterales se rechazan  $\rightarrow \varepsilon_1 < \mu_N - \mu_R < \varepsilon_2 \rightarrow$  equivalencia

[Nótese que  $H_{01}$  y  $H_{02}$  no pueden ser simultáneamente ciertas, ergo sus riesgos  $\alpha$  correspondientes no se suman. Por convención, se acepta como nivel de significación  $P_1, P_2$  observados.]

**Ejemplo:** Equivalencia Clínica:

Hipotensor: equivalencia si  $|diferencia| < 10$  mmHg.

Resultados:  $n = 20$ ; datos apareados:  $\bar{X}_D = 3$  y  $S_D = 10$

Solución mediante **Prueba Doblemente Unilateral** ( $\alpha = 0.05$ )

$$\begin{cases} H_{01} : \mu_R - \mu_N \leq -10 \\ H_{11} : \mu_R - \mu_N > -10 \end{cases} \quad t_1 = (3 - (-10)) / (10 / \sqrt{20}) = 5.814 \rightarrow P < 0.001$$

Luego la diferencia entre las medias está por encima de -10

$$\begin{cases} H_{02} : \mu_R - \mu_N \geq 10 \\ H_{12} : \mu_R - \mu_N < 10 \end{cases} \quad t_2 = (3 - 10) / (10 / \sqrt{20}) = -3.130 \rightarrow P \geq 0.003$$

Luego la diferencia entre las medias está por debajo de 10.

Por tanto, con un nivel de significación  $P \geq 0.003$ , se ha establecido que la diferencia entre las medias de ambos hipotensores está entre -10 y +10.

**Decisión basada en el intervalo de confianza**

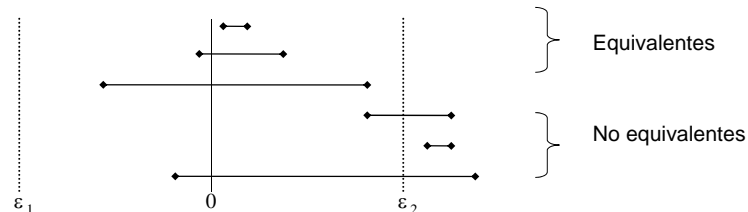
Pueden utilizarse los IC clásicos,

simétricos alrededor de la estimación puntual ( $\bar{X}_1 - \bar{X}_2$ )

del parámetro a estimar  $\mu_N - \mu_R$ ,

como criterio de decisión en el estudio de equivalencia:

Se concluye equivalencia si el IC está *enteramente comprendido* entre  $\varepsilon_1$  y  $\varepsilon_2$



Este criterio es idéntico al PDU si el IC se calcula con una confianza  $1 - 2\alpha$ .

Es decir: 0.90 si  $\alpha = 0.05$ .

Aunque la estimación por IC se realiza con una confianza  $1 - 2\alpha = 0.90$ , el *criterio de decisión* basado en dicho IC tendrá un riesgo de errores tipo I  $\alpha = 0.05$ .

**Ejemplo (continuación):**

Solución mediante el intervalo de confianza (90%).

$$(\mu_R - \mu_N) \in (\bar{X}_R - \bar{X}_N) \pm t_{1-2\alpha} \cdot S/\sqrt{n} \rightarrow$$

$$(\mu_R - \mu_N) \in (-0'866, 6'866)$$

La diferencia entre las medias de ambos hipotensores está entre  
 -0,866 (gana el de referencia por 0,866mmHg) y  
 +6'866 (gana el nuevo por 6'866mmHg).

Como la diferencia entre ambos fármacos está entre los límites -10 y +10:  
 se ha establecido equivalencia.

**Ejemplo:**

Un nuevo (N) antiinflamatorio tiene una tolerabilidad superior a cierto  
 producto clásico de referencia (R).

Interesa poder demostrar que son parecidos sus niveles de eficacia, que se  
 mide por el % de casos en los que desaparece el dolor a los 30'.  
 Ambos fármacos serán declarados como equivalentes en eficacia si las  
 proporciones de desaparición del dolor no difieren en más de un 8%  
 ( $\varepsilon$ ).

A) El intervalo de confianza (90%) de la diferencia de ambas proporciones  
 va entre -6% y +3%.

Dado que no alcanza los límites de la no equivalencia, se puede  
 rechazar esta última ( $\alpha=0'05$ )

B) La PDU rechaza ambas hipótesis nulas de no equivalencia ( $P_1=0'02$  y  
 $P_2=0'0003$ )

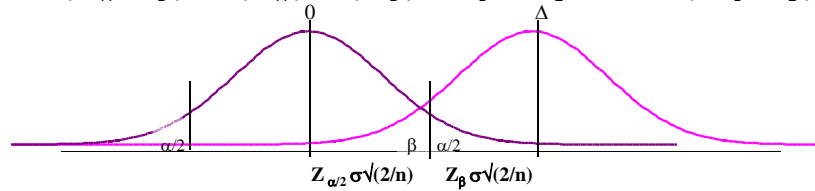
Se concluye equivalencia con un nivel de significación  $P=0'02$ .

### Tamaño muestral.

Recuerde: en hipótesis de diferencias sean:

- El contraste de hipótesis:  $H_0 : \mu_A - \mu_B = 0$   
 $H_1 : \mu_A - \mu_B = \Delta$
- Un riesgo alfa, uni o bilateral (usualmente  $\alpha=0.05$ , bilateral),
- La potencia deseada (usualmente  $1-\beta=0.8$ ),
- La desviación típica del término de error ( $\sigma$ ), y
- La magnitud del efecto que se desea establecer  $\Delta$ .

$$V(x_A^- - x_B^-) = V(x_A^-) + V(x_B^-) = \sigma^2/n_1 + \sigma^2/n_2 = 2\sigma^2/n \quad (\text{si } n_1 = n_2)$$



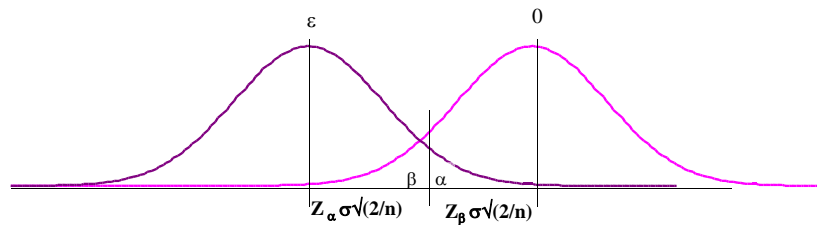
$$\Delta = Z_{\alpha/2} \sigma \sqrt{2/n} + Z_{\beta} \sigma \sqrt{2/n}$$

$$\text{de donde el tamaño por grupo } n = [2 \sigma^2 (Z_{\alpha/2} + Z_{\beta})^2] / \Delta^2$$

En el caso de hipótesis de Equivalencia:

$$\text{Ahora el contraste de hipótesis es : } \begin{cases} H_0 : \mu_A - \mu_B = \varepsilon \\ H_1 : \mu_A - \mu_B = 0 \end{cases}$$

Tan sólo intercambia  $H_0$  y  $H_1$



Como el planteamiento es unilateral

$$n = [2 \sigma^2 (Z_{\alpha} + Z_{\beta})^2] / \varepsilon^2$$

¡MISMA FÓRMULA!

Pero,

1) Ahora  $\varepsilon$  representa la diferencia irrelevante, necesariamente inferior a la relevante  $\Delta \rightarrow$  Mayor número de casos

2) Podría ser que los productos no fueran idénticos, pero aún así puedan ser equivalentes: la distancia entre  $H_0$  y  $H_1$  sería menor a  $\varepsilon$ . Si, por ejemplo, la auténtica diferencia fuera  $\Delta$ , el contraste sería :

$$\begin{cases} H_0: \mu_A - \mu_B = \varepsilon \\ H_1: \mu_A - \mu_B = D \end{cases}$$

Conduciendo a:

$$n = [2\sigma^2(z_\alpha + z_\beta)^2] / (\varepsilon - D)^2$$

$\rightarrow$  Mayor número de casos

## Tipos de equivalencia

**a) Poblacional:** equivalencia en media.

Implica que el valor esperado de la administración del fármaco sea el mismo para ambos productos: PRESCRIBIBILIDAD

**b) Individual:**

Si no existe interacción tratamiento-caso (o en otras palabras si es un efecto fijo, constante), el valor esperado para un caso concreto es el mismo si se le cambia el tratamiento: INTERCAMBIABILIDAD

## Sensibilidad del ensayo

**Sensibilidad de un ensayo:** capacidad de un ensayo clínico concreto para distinguir entre un tratamiento eficaz y un tratamiento ineficaz o menos eficaz.

Ensayos de superioridad: si se detectan diferencias

- el estudio tiene sensibilidad (capaz de ver diferencias)
- puede concluirse a favor de la eficacia

Ensayos de no-inferioridad: si no se detectan estas diferencias

- ¿el estudio tiene sensibilidad?
- no puede concluirse a favor de la equivalencia

La sensibilidad puede ser clara si:

existen ensayos controlados con placebo de resultados consistentes  
la evolución de los casos tratados y no tratados es muy distinta  
(p.ej., en enfermedades infecciosas)

## Factores que pueden reducir la sensibilidad del ensayo:

### Diseño del ensayo

- Cambios en la población en estudio (criterios de selección):  
severidad de la patología, enfermedades concomitantes, métodos diagnósticos
- Cambios en la dosis y pauta del tratamiento control
- Cambios en las variables de eficacia y su momento de evaluación
- Periodos de lavado pre-inclusión

### Desarrollo del ensayo

- Bajo cumplimiento con la medicación
- Baja respuesta de los pacientes a los tratamientos
- Uso de tratamientos concomitantes prohibidos
- Pacientes que tiendan a mejorar espontáneamente
- Criterios diagnósticos mal aplicados (pacientes sin la patología)
- Evaluación sesgada debida al conocimiento de que todos los pacientes reciben algún tratamiento activo