# TEMA 2: ENSAYO CLÍNICO

# DISEÑOS CON INTERCAMBIO (*CROSS-OVER*)

## 1.- VARIANZA INTRA Y ENTRE

# Si el proceso de medida es razonable, siempre cabrá esperar σ2τ >> σ2I

# 3.- UN SÍNDROME RARO.

a) Hipótesis: H0: μN =μP versus H1: μN ≠μP

H0: Igual media del número de crisis bajo ambos tratamientos.

(Efecto directo del tratamiento).

Premisas: Ausencia de efecto tardío (o *carry-over*) del tratamiento.

Se sostiene por el periodo de lavado.

+ las paramétricas habituales

(Muestra aleatoria, Normalidad, homocedasticidad)

Valor del estadístico y distribución teórica:

Spooled =[ (n1-1)S12 + (n2-1)S22 ] / (n1 +n2 –2) ≈ 13’492

T = (-1 –12) / [13’49√(1/10 + 1/10)] ≈ -2’15

-2’15 < - 2’101 = t18,0’025 → Se rechaza H0 con riesgo α

Nicardipino reduce, respecto a placebo, el nº de crisis.

b) IC(α,0’95) = ½ (-1-12) ± ½ · 2’101 · 13’49 · √(1/10 + 1/10) =[-6’5 ± 6’34] = [-12’84,-0’16]

Con una confianza del 95%, la reducción del número de crisis es algún valor entre -12’84 y -0’16. Dado que excluye el valor cero (no efecto), coincide con el resultado anterior.

NOTA: El enunciado permite no tener que establecer la ausencia de efecto residual, mediante la comparación de la variable suma, al considerarse suficiente el efecto de lavado. [Recuérdense las críticas sobre poner a prueba esta premisa].

c) El 80% de 4 crisis sería reducir 3’2 crisis. Este es un resultado compatible con el intervalo anterior, pero también lo son valores entre 0’16 y 3’2 crisis, que indicarían que el nuevo tratamiento no alcanza el efecto deseado

Es decir, dado que el intervalo de confianza incluye -3’2, no se puede afirmar, con la confianza especificada, que el efecto se sitúe a la izquierda de -3’2.

4.- ASPIRINA A GRANDES DOSIS

Efecto **directo** del tratamiento (se basa en las restas):

t = [1.4875-(-0.3625)]/1.4255√(1/8+1/8) = 2.6

[donde 1.4255=√(1.4904²+1.3575²)/2]

t gdl,α/2 = t 14,0,025 = 2.145

Como 2.6 > 2.145, el efecto directo es estadísticamente significativo (α=0.05) con intervalo de confianza (recuérdese que debe dividirse por 2):

½[1.4875-(-0.3625)] ± ½2.145[1.4255√(1/8+1/8)] →

→ 0.161 ≤ ΓA - ΓB ≤ 1.689

Premisas: No efecto tardío, residual o *carry-over,*

+ las paramétricas:

Normalidad de la variable diferencia,

igualdad de varianza de esta variable en ambas secuencias,

independencia de estas diferencias.

Efecto **residual** del tratamiento (se basa en las sumas):

t = (7.2625-6.1375)]/2.8653√(1/8+1/8) = 0.79

[donde 2.8653=√(3.1645²+2.5309²)/2]

Como 0.79 < 2.145, no es estadísticamente significativo (nada se opone a aceptar la igualdad de los efectos residuales o tardíos del tratamiento).

Efecto **período**:

t = [1.4875+(-0.3625)]/1.4255√(1/8+1/8) = 1.58

Que, con 14 gdl, no es estadísticamente significativo.

5.- UN EJEMPLO DE BOLTON (1994).

X– 2 – X– 1 = - 0,0542 + 0’4937 = 0’4395

→ estimador de μln(N/C) = ½ 0’4395 = 0’2197

S = √ {[S21 (n1 –1)+ S21 (n2 –1)]/(n1+n2 –2)} = √{(0’22922 + 0,35682)/2} = 0,29985

→ S(est) = 0’14992

ln(0’8) = -0’2231 ln(1’25) = 0’2231 t(10,0’95) = 1’8125

a) Mediante Intervalo de Confianza:

I.C.(90%) = est ± t(10,0’95) · S(est)√(2/n) = 0’2197 ± 1’8125 · 0’14992√(2/6)

L.I. (ln(C/N)) = 0,0628 → L.I. (C/N) = 1,0649

L.S. (ln(C/N)) = 0,3766 → L.I. (C/N) = 1,4573

Con una confianza del 90%, la auténtica media de la diferencia de los logaritmos de la concentración del preparado clásico menos el nuevo se encuentra entre 0’0628 y 0’3766, fuera del intervalo de bioequivalencia (-0’2231, 0’2231).

[O bien: Con una confianza del 90%, la auténtica media geométrica del cociente entre la concentración del preparado clásico y la del nuevo se encuentra entre 1’0649 y 1’4573, fuera del intervalo de bioequivalencia (0’8, 1’25)].

→ Los preparados no son bioequivalentes

b) Mediante Prueba Doblemente Unilateral:

H01 : μln(N/C) ≤ -0’2231 frente a H11 : μln(N/C) > -0’2231

t1 = (0’2197-(-0’2231))/( 0’1499√(2/6)) ≅ 0’4429/0’0866 ≅ 5’116

→P<0’001

Luego la diferencia entre las medias está por encima de –0’2231

H02 : μln(N/C) ≥ 0’2231 frente a H12 : μln(N/C) < 0’2231

t2 = (0’2197-(0’2231))/( 0’1499√(2/6)) ≅-0’0034/0’0866≅-0’0395

→P≅0’485

Luego no podemos rechazar que la diferencia esté por encima de 0’2231

No hemos conseguido demostrar que

la diferencia esté comprendida entre –0’2231 y +0’2231.

→ No hemos demostrado la bioequivalencia.

6.- REANALIZANDO EL EJEMPLO DE BOLTON (1994).

No es un buen sistema para establecer la ausencia de efecto tardío o residual:

1. Tiene menos potencia que el estudio del efecto directo del tratamiento, ya que no suprime la variabilidad interindividual,
2. Pretende establecer la hipótesis nula, pero la ausencia de significación estadística no implica el establecimiento de la hipótesis nula (y menos aún si ya sabemos que falta potencia). Debería acompañarse de cálculos de potencia (lo que implica establecer un valor de la hipótesis alternativa para calcular esta potencia), o intervalos de confianza, pruebas de equivalencia,...
3. No tiene sentido, desde una perspectiva confirmatoria, ir haciendo pruebas bajo diferentes distribuciones. Debería estudiarse bien la distribución mediante argumentación teórica sobre el origen de la variable y mediante gráficos de normalidad. Sin embargo, “consuela” que todas hubieran conducido a la misma conclusión. Por supuesto, no tiene sentido realizar ajustes por multiplicidad, ya que todavía “castigaríamos” más la falta de potencia.
4. Sorprende que la prueba de los rangos de Mann-Whitney muestre diferente resultado tras la transformación logarítmica, ya que es monotónica, no altera el orden y lo rangos serían los mismos. Sucede porque la prueba se aplica a la suma de dos variables y la transformación se realiza antes de la suma.

Una vez más, conviene garantizar la premisa de no efecto tardío por otros caminos.

7.- EL EJEMPLO DEL SENN (1993)

Estamos en el caso de conocer σ2d, por lo que:

n = [ σ2d (Z α/2 +Z β)²] / Δ ² =

= [452 (1’96 + 0’84)2 ] / 302 ≅

≅ 17’64

que redondearíamos al entero superior 18.

Nótese que como al final suele utilizarse la estimación la varianza muestral, en realidad se utiliza una ‘t’ de Student, por lo que, utilizando los grados de libertad obtenidos del cálculo anterior:

n = [ σ2d (t 17, α/2 +t 17,β)²] / Δ ² ≅ [452 (2’1098 + 0’8633)2 ] / 302 ≅19’8884

lo que utilizado iterativamente

n = [ σ2d (t 19, α/2 +t 19,β)²] / Δ ² ≅ [452 (2’0930 + 0’8610)2 ] / 302 ≅ 19’6334

nos conduce a proponer una muestra de 20 casos que, al ser par, permite diseñar dos secuencias AB, BA con el mismo número de casos.

[NOTA: La inclusión del efecto periodo en el modelo conllevaría únicamente la pérdida de un grado de libertad, por lo que este cálculo puede considerarse válido. Obsérvese que repetir el proceso anterior con 18 grados de libertad conduciría a un resultado intermedio entre los de 19 y 17 grados de libertad].

8.- CASOS NECESARIOS PARA UN ESTUDIO CON INTERCAMBIO.

Ahora tenemos la variabilidad intraindividual:

n = [ 2 σ2I (Z α/2 +Z β)²] / Δ ² = [2·202 (1’96 + 0’84)2 ] / 152 ≅ 27’88

Igualmente, podemos calcular la varianza de la diferencia entre ambos periodos:

V(D) = 2 σI2 = 2·202 = 800

* σD = √800 ≅ 28’3

n = [ σ2d (Z α/2 +Z β)²] / Δ ² = [800 (1’96 + 0’84)2 ] / 152 ≅ 27’88

Se usarían 28 casos, 14 en cada secuencia.

9.- TAMAÑO DISTINGUIENDO ENTRE VARIABILIDAD ENTRE E INTRA.

Paralelo: σ² = σ²Γ+ σ²I = 200

n = [ 2 σ² (Z α/2 +Z β)²] / Δ ² = 2 ·200 (1'96+0'84)²/10² = 31'36 -> 32

Implica 32 casos por grupo = 64 casos en total.

Con intercambio: σ²I = 100

n = [ 2 σ²I (Z α/2 +Z β)²] / Δ ² = 2· 100 (1'96+0'84)²/10² = 15'62 -> 16

Se necesitan 16 casos en total.

10.- EFECTO PERIODO

Los resultados obtenidos durante el primer periodo tienen una media diferente que los del segundo periodo. Varios fenómenos pueden explicar este hecho: una enfermedad que no es estable a lo largo del tiempo (por curación, mejoría, empeoramiento,…); una alteración del procedimiento de medida (diferentes condiciones ambientales, pérdida del calibrado del aparato,…); diferente evaluador en ambos periodos, …

Mientras el diseño sea balanceado, y a diferencia de lo que sucede con el efecto residual, un efecto periodo no está confundido con el efecto directo, que se puede seguir estimando:

## E(d1) = (αA - αB)+ (Π1 - Π2) +ΟA

## E(d1) = (αB - αA)+ (Π1 - Π2) +ΟB

## E(d1-d2) = 2(αA - αB) +(ΟA- ΟB)

[Otro tema aparte es si cabe esperar el mismo efecto del tratamiento en unas condiciones diferentes a las basales.]

11.- DISEÑO CON INTERCAMBIO PARA COMPARAR A Y B (1)

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Secuencia | AB | | | | Secuencia | BA | | | |
| Periodo | 1 | 2 |  |  | Periodo | 2 | 1 |  |  |
| Caso | A | B | d | S | Caso | A | B | D | S |
| 1 | 12 | 8 | 4 | 20 | 5 | 12 | 9 | 3 | 21 |
| 2 | 11 | 7 | 4 | 18 | 6 | 10 | 7 | 3 | 17 |
| 3 | 10 | 7 | 3 | 17 | 7 | 10 | 6 | 4 | 16 |
| 4 | 11 | 8 | 3 | 19 | 8 | 11 | 8 | 3 | 19 |
| Media | | | 3,500 | 18,500 | Media | | | 3,250 | 18,250 |
| Desviación típica | | | 0,577 | 1,291 | Desviación típica | | | 0,500 | 2,217 |

a) S2d, pooled = (3·0’5772 + 3·0’5002)/ 6 ≈ 0’2917 ≈ 0’542

## *Tablas: T6,0’975 ≈ 2’447*

IC95% ≈ ½ (3’5 +3’25) ± ½ 2’447·0’54 √(¼+¼) ≈ 3’375 ± 0’4671 ≈ [2’91, 3’84]

b) S2S, pooled = (3·1’2912 + 3·2’2172)/ 6 ≈ 3’2917 ≈ 1’8142

V(d) = 2 σI2 → Sε2 ≈ 0’2917 / 2 ≈ 0’1458

V(S) =2 σI2 + 4 σΤ2 → SΤ2≈ (3’2917-2·0’1458)/4 ≈ ¾ =0’75

[Nótese que la varianza entre, SΤ2, es sensiblemente superior (unas cinco veces) a la intra, SI2, lo que parece razonable. La fiabilidad de la medida rondaría los 5/6. ]

12.- ENEURESIS.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variable Suma | | | | |
| Orden | N | media | S | Sconjunta |
| F P | 17 | 13’41 | 7’323 | 6’472 |
| P F | 12 | 16’58 | 4’981 |

1. IC95%(μF-μP) = (13’41-16’58) ± 2’052·6’472√(1/17+1/12) =

= -3’17 ± 2’052 · 2’44 =

= -3’17 ± 5’007 =

= [ -8’18 , +1’835 ]

Dado que el intervalo contiene el valor 0, que implica no efecto tardío (o, hablando en propiedad, efectos tardíos de igual magnitud), tiene sentido estimar, mediante este procedimiento, el efecto directo del tratamiento.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variable Diferencia | | | | |
| Orden | n | media | S | Sconjunta |
| F P | 17 | 2’824 | 3’468 | 3’281 |
| P F | 12 | -1’25 | 2’989 |

t27,0’0975 = 2’052

IC95%(μF-μP) = ½ [ (2’824-(-1’25)) ± 2’052·3’281√(1/17+1/12) ] =

= ½ [ 4’074 ± 2’052 · 1’237 ] =

= ½ [ 4’074 ± 2’ 537 ] =

= ½ [ 1’535 , 6’612 ] =

= [ 0’768, 3’306 ]

Dado que el intervalo no contiene el valor 0, que implicaría no diferencia del efecto directo del tratamiento, rechazamos la hipótesis de no efecto de este tratamiento farmacológico y decimos que desciende más que placebo el número de micciones nocturnas en 14 días entre 0’768 y 3’306 noches.

c) Se trata de un efecto tardío, estimable por la pregunta (a), que, de existir, sesgaría la estimación del efecto directo realizada en la pregunta (b).

NOTA: No se trata de un efecto periodo, ya que no se manifiesta en todos los casos en el segundo periodo, sino sólo en los casos en los que hubo efecto directo en el primer periodo. El paso del tiempo, la maduración del niño,… podrían conllevar un efecto periodo que no sesgaría la estimación realizada en (b), aunque no sería “elegante” su presencia.

13.- MÉTODO DOCENTE (prueba final 28/05/03)

# 1) S2D,pooled = (0’62 + 0’72)/2 = 0’85/2 = 0’425 = 0’6522

# Tablas T18, 0’975 *≈ 2’10*

# *IC95% = ½ (1’9 + 1’2)* ± ½ 2’10 · 0’652 √(1/10 + 1/10) = 1’55 ± *0’30616 = [1’25, 1’85]*

# 4) V(dif) = V(A) + V(B) - 2cov(A,B) = 0’4225

# V(suma) = V(A) + V(B) + 2cov(A,B) = 4’2025

# Suma 4V(A) = 4’625

# Resta - 4 cov(A,B) = -3’78

# r=SAB/SA·SB = SAB/SA2 = 3’78/4’625 *≈ 0’82*

14.- DISENYO CON INTERCAMBIO PARA COMPARAR A Y B (2) (tercer examen parcial 22/12/06)

a)Puntual: τ = (2’5 +2’75)/2 = 2’625

Sd =(3·0’5772+3·0’52)/6 ≈ 0’2914 ≈ 0’53982

IC95%T = 2’625 ± ½ [2’477·0’5398·√(¼ + ¼ ) ≈2’625 ± 0’467≈ [2’158,3’092]

b) Puntual: π = (2’5 -2’75)/2= 0’125

Ss =(3·0’5772+3·0’52)/6 ≈ 0’2914 ≈ 0’53982

IC95%π = 0’125 ± ½ [2’477·0’5398·√(¼ + ¼ ) ≈0’125 ± 0’467≈ [-0’364,0’592]

c) Puntual: O = (5’5 -4’75) = 0’75

Sd =(3·1’2912+3·1’7082)/6 ≈ 2’2919 ≈ 1’51392

IC95%O = 0’75 ± 2’477·1’5139·√(¼ + ¼ ) ≈ 0’75 ± 2’6194≈ [-1’8694,3’3694]

***d)*** intra = S2pool(periodo)/2 ≈ 0’146

entra = S2pool(tardio)-S2pool(periodo)/4 ≈ 0’5 total = intra + entre ≈ 0’646

fiabilidad = entre / total ≈ 77’42%

15.- MÁS DE JAQUECAS

1. Ponga a prueba y estime por intervalo el efecto directo del tratamiento.

S2P- = √(142·9+162·9/18) = √(142+162/2) ≈ 15.033

t = [24-(-2)] / [15.033√(1/10 +1/10)] ≈ 26 / 6.7231 ≈ 3.8673 > 2.10=t18,0.025

Se rechaza H0 y, por tanto, se ha logrado demostrar el efecto directo.

IC95%τ = ½ [24-(-2)] ± ½ 2.10·15.033√(1/10 +1/10) ≈ 13 ± 7.063 ≈ [5.94, 20.06]

1. Ponga a prueba y estime por intervalo el efecto tardío del tratamiento.

S2P+ = √(292·9+212·9/18) = √(292+212/2) ≈ 25.3179

t = [87-76)] / [25.32√(1/10 +1/10)] ≈ 11/11.3225 ≈ 0.9715 < 2.10=t18,0.025

No se puede rechazar H0 y, por tanto, no se ha logrado demostrar el efecto tardío.

IC95%τ = [87-76] ± 2.10·25.32√(1/10 +1/10) ≈ 11 ± 23.7773 ≈ [-12.77, 34.77]

1. Indique cómo estimaría por intervalo el efecto período.

S2P- = √(142·9+162·9/18) = √(142+162/2) ≈ 15.033

t = [24+(-2)] / [15.033√(1/10 +1/10)] ≈ 22 / 6.7231 ≈ 3.2723 > 2.10=t18,0.025

Se rechaza H0 y, por tanto, se ha logrado demostrar el efecto periodo.

IC95%τ = ½ [24+(-2)] ± ½ 2.10·15.033√(1/10 +1/10) ≈ 11 ± 7.063 ≈ [3.94, 18.06]

1. Estime puntualmente los valores de σ2E y σ2I

σ2I (intra) = S2P- / 2 = 15.03332 / 2 ≈ 113.00

σ2E (entre) = (S2P+ - S2P- ) / 4 = (25.31792-15.03332)/4 ≈ 103.75

1. Calcule la correlación entre los resultados alcanzados con cada tratamiento

# BIO-EQUIVALENCIA

# 1.- UN HIPOTENSOR CON MENOS EFECTOS SECUNDARIOS.

Solución mediante la Prueba Doblemente Unilateral (α=0’05)

H01 : μR - μN ≤ -10

H11 : μR - μN > -10 t1 = (3-(-10))/(10/√20) =5’814 → P<0’001

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

H02 : μR - μN ≥ 10

H12 : μR - μN < 10 t2 = (3- 10 )/(10/√20) =-3’130 → P≅0’003

Luego la diferencia entre las medias de ambos hipotensores está por debajo de 10

Luego, con un nivel de significación P≅0’003, hemos establecido que la diferencia entre las medias de ambos hipotensores está entre –10 y +10.

b) Solución mediante el intervalo de confianza (90%).

(μR - μN) ε(XR – XN) ± t1-2α ·S/√n → (μR - μN) ε (-0,866 , 6’866)

Luego la diferencia entre las medias de ambos hipotensores está entre –0,866 (el de referencia consigue presiones mas bajas en media: gana por 0,866 mmHg) y +6’866 (el nuevo consigue presiones más bajas: gana por 6’866mmHg). Luego la diferencia entre ambos fármacos está entre los límites –10 y +10: se ha establecido equivalencia.

2.- MISMO HIPOTENSOR PERO MENOS CASOS.

a ) Solución mediante la Prueba Doblemente Unilateral (α=0’05)

H01 : μR - μN ≤ -10

H11 : μR - μN > -10 t1 = (3-(-10))/(10/√5) =2’907 → P≅0’022

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

H02 : μR - μN ≥ 10

H12 : μR - μN < 10 t2 = (3- 10 )/(10/√5) =-1’565 → P≅0’096

Luego no podemos rechazar la posibilidad de que la diferencia de las medias de ambos hipotensores esté por encima de 10

No hemos conseguido demostrar que la diferencia entre las medias de ambos hipotensores esté comprendida entre –10 y +10.

b) Solución mediante el intervalo de confianza (90%).

(μR - μN) ε(XR – XN) ± t1-2α ·S/√n → (μR - μN) ε (-6’534 , 12’534)

No hemos conseguido demostrar que la diferencia entre las medias de ambos hipotensores esté comprendida entre –10 y +10.

3.- MISMO HIPOTENSOR PERO DIFERENTE HIPÓTESIS.

Ahora se trata de la hipótesis de “al menos equivalencia”. Se desea que el nuevo hipotensor tenga consiga presiones que no sean superiores a la del clásico en más de 10 mmHg.

a) Solución mediante Prueba Unilateral (α=0’05)

H01 : μR - μN ≤ -10

H11 : μR - μN > -10 t1 = (3-(-10))/(10/√5) =2’907 → P≅0’022

Luego la diferencia entre las medias de ambos hipotensores está por encima de –10.

R no gana a N por una diferencia de 10 mmHg. Hemos demostrado que el nuevo hipotensor es, al menos, equivalente al de referencia.

b) Solución mediante el intervalo de confianza unilateral (95%).

(μR - μN) ε [ (XR – XN) - t1-2α ·S/√n , ∞ ] → (μN - μR) ε (-6,534 , ∞)

Como mucho, el fármaco de referencia baja la presión arterial 6,534mmHg más que el fármaco nuevo.

4.- GENÉRICOS.

a) ¿Se cumplen los criterios de Bioequivalencia?

Solución mediante la Prueba Doblemente Unilateral (α=0’05)

H01 : μLR-LN ≤ ln(0’8)

H11 : μLR-LN > ln(0’8) t1 = (0’1-(-0’223))/(0’2/√12) =5’597 → P<0’001

Luego la media de la diferencia entre los logaritmos está por encima de –0’223

→ la media geométrica del cociente está por encima de 0’8

H02 : μLR-LN ≥ ln(1’25)

H12 : μLR-LN < ln(1’25) t2 = (0’1- 0’223 )/(0’2/√12) =-2’133 → P≅0’028

Luego la media de la diferencia entre los logaritmos está por debajo de 0’223

→ la media geométrica del cociente está por debajo de 1’25

Hemos demostrado (P=0’028)

que la media geométrica del cociente está entre 0’8 y 1’25.

Solución mediante el intervalo de confianza (90%).

μLR-LN εXLR-LN ± t1-2α ·SLR-LN/√n → μLR-LN ε (-0,004 , 0’204)

Luego la media de la diferencia entre los logaritmos está entre –0’004 y 0’204

→ la media geométrica del cociente está entre 0’996 y 1’226

es decir, dentro del intervalo (0’8, 1’25) o de bioequivalencia

b) La premisa de normalidad aplica a la variable diferencia de los logaritmos. Si las concentraciones en sangre tuvieran una distribución lognormal, dicha diferencia sería normal al ser una diferencia de normales. [También podría ser normal la diferencia de los logaritmos sin que fueran normales los logaritmos de las concentraciones.]

c) Nótese que si los niveles en sangre de un genérico tienen mayor dispersión que el fármaco de referencia, ello implica que los niveles individuales serán más extremos, pudiendo alcanzar la ineficacia –por el lado inferior- o la toxicidad –por el lado superior.

## 

## 5.- COMPARACIÓN DE PRUEBAS, GRÁFICOS

a) Límites de significación para la prueba doblemente Unilateral

HO : N - R  1 ó N - R  2  No equivalencia

HA : 1 < N - R < 2   Equivalencia

s(2/n)

25

##### 20

15

No se ve equivalencia

No se ve equivalencia

10

5

Equivalencia

0

-20 -10 0 10 20 x–1 - x– 2

b) Límites de significación para la prueba clásica de diferencias

HO : N - R = 0  Igualdad

HA : N - R ≠ 0  No Igualdad 🡪 diferencias (alternativa que quiero demostrar)

s(2/n)

25

##### 20

15

10

No se ven diferencias

5

Diferencias

Diferencias

0

-20 -10 0 10 20 x–1 - x– 2

a) Límites de significación para ambas pruebas

s(2/n)

25

##### 20

E

15

D

10

5

C

A

B

0

-20 -10 0 10 20 x–1 - x– 2

|  |  |  |
| --- | --- | --- |
|  | Prueba clásica de diferencias | Prueba doblemente unilateral |
| A | Nada se opone a aceptar la igualdad de medias  NO HAY CONCLUSIÓN | Se rechazan ambos límites de equivalencia al riesgo α. SE CONCLUYE EQUIVALENCIA:  -20 < μ1-μ2 < 20 |
| B | Se rechaza la igualdad de medias con riesgo α SE CONCLUYEN DIFERENCIAS 0 < μ1-μ2 | Se rechazan ambos límites de equivalencia al riesgo α. SE CONCLUYE EQUIVALENCIA  -20 < μ1-μ2 < 20 |
| C | Se rechaza la igualdad de medias con riesgo α SE CONCLUYEN DIFERENCIAS 0 < μ1-μ2 | Nada se opone a aceptar la no equivalencia  para el límite superior  NO HAY CONCLUSIÓN |
| D | Nada se opone a aceptar la igualdad de medias  NO HAY CONCLUSIÓN | Nada se opone a aceptar la no equivalencia  para el límite superior  NO HAY CONCLUSIÓN |
| E | Nada se opone a aceptar la igualdad de medias  NO HAY CONCLUSIÓN | Nada se opone a aceptar la no equivalencia  para ambos límites  NO HAY CONCLUSIÓN |

La zona más problemática es la B, en la que se puede demostrar tanto que los productos son diferentes como que son equivalentes. Ello se explica porque la región de equivalencia incluye las diferencias no relevantes.

Las zonas D y E, son algo contradictorios, pero se explica por el elevado error típico de estimación, que las hace incapaces de demostrar nada, por lo que no logran ni establecer equivalencia ni demostrar diferencias.

Las zonas A y C no son problemáticas.

**META-ANÁLISIS**

1.- ASPIRINA

OR ln(OR) V(Ln(OR)) LI(Ln(OR)) LS(Ln(OR)) LI(OR) LS(OR)

1 1,3894 0,3289 0,039 -0,058 0,715 0,944 2,045

2 1,4689 0,3845 0,041 -0,013 0,782 0,987 2,186

3 1,2455 0,2196 0,020 -0,061 0,500 0,941 1,649

4 1,2488 0,2222 0,065 -0,277 0,721 0,758 2,057

5 1,2529 0,2255 0,035 -0,142 0,593 0,867 1,810

6 0,8828 -0,1246 0,010 -0,317 0,068 0,728 1,070

7 1,1174 0,1110 0,002 0,035 0,187 1,036 1,206

EC Wc Wc\*Yc Wc(Yc-Y)^2 Wc^2 Wc\* Wc\*Yc

1 25,710 8,456 1,246 660,995 20,605 6,777

2 24,291 9,341 1,847 590,060 19,683 7,569

3 48,801 10,715 0,599 2381,509 33,191 7,287

4 15,441 3,431 0,199 238,410 13,441 2,987

5 28,409 6,405 0,387 807,093 22,303 5,029

6 103,985 -12,960 5,666 10812,969 51,937 -6,473

7 663,923 73,673 0,003 440794,250 89,738 9,958

Suma 910,560 99,061 456285,285 250,898 33,133

Q= 9,946 α=0,050 W– = 130,080 U= 409,457 Ef Fijos D=0,000

Zα/2=1,960 SW2= 56306,590 (Q-(C-1))/U= 0,010 Ef Aleat D= 0,010

Y V(Y) LI(Y) LS(Y)

Efectos Fijos 0,109 0,001 0,044 0,174

Efectos Aleat 0,132 0,004 0,008 0,256

OR LI(OR) LS(OR)

Efectos Fijos 1,115 1,045 1,190

Efectos Aleat 1,141 1,008 1,291

2.- UN META-ANALISIS DE ANTIHISTAMÍNICOS.

El estudio individualizado de los 9 EC muestra que sólo el 7 y el 9 tienen intervalos que excluyen el valor de no efecto (0) y que, por lo tanto, sólo estos dos EC hubieran permitido establecer efecto.

EC pooled S Yc V(Yc) L.I(Yc) L.S.(Yc)

1 0,815 0,57 0,15 -0,20 1,33

2 0,574 0,21 0,02 -0,03 0,45

3 0,759 0,20 0,03 -0,15 0,55

4 0,962 0,01 0,09 -0,58 0,60

5 1,853 0,24 0,13 -0,47 0,94

6 1,640 -0,17 0,05 -0,61 0,27

7 1,537 0,93 0,10 0,33 1,54

8 1,800 0,59 0,20 -0,29 1,47

9 0,862 0,26 0,01 0,06 0,46

EC Wc Wc\*Yc Wc(Yc-Y)^2 Wc^2 VAR(aleat) Wdc Wdc\*Yc

1 6,52 3,69 0,70 42,49 0,17 6,02 3,41

2 65,94 14,01 0,04 4348,00 0,03 35,99 7,65

3 31,75 6,36 0,04 1007,94 0,04 22,67 4,54

4 11,00 0,13 0,56 121,00 0,10 9,66 0,11

5 7,74 1,83 0,00 59,94 0,14 7,05 1,67

6 19,99 -3,43 3,35 399,50 0,06 15,96 -2,74

7 10,50 9,80 5,08 110,25 0,11 9,27 8,66

8 4,95 2,92 0,61 24,53 0,21 4,66 2,75

9 96,75 25,36 0,06 9360,44 0,02 43,56 11,42

Suma 255,14 60,67 10,45 15474,08 154,85 37,46

Q=10,454 α=0,050 W– =28,349 U=194,487 Ef Fijos D=0,000

P=0,235 Zα/2=1,960 SW2=1030,164 (Q-(C-1))/U=0,013 Ef Aleat D=0,013

Y SE(Y) LI(Y) LS(Y)

Efectos Fijos 0,238 0,063 0,115 0,361

Efectos Aleat 0,242 0,080 0,084 0,399

Premisas: La inspección visual apoya la homocedasticidad. Dado lo poco que sabemos sobre la distribución de la variable respuesta “porcentaje de reducción de sintomatología”, los pequeños tamaños de los EC 8, 5, 7 y 4 podrían desaconsejar el uso de la distribución normal en el cálculo de los intervalos de confianza individuales (pero no el global)

Si se parte del principio de que el efecto del tratamiento es el mismo en todos los estudios, la prueba de Homogeneidad no arroja resultados significativos, y la conclusión es que los antihistamínicos bajan más que el placebo la intensidad de la sintomatología basal. Este descenso, común y único en todos los estudios es, con una confianza del 95%, alguno de los valores comprendidos entre 11’5% y 36’1%.

Si se acepta cierta variabilidad del efecto de los tratamientos debido a las diferencias metodológicas entre los diferentes estudios, la esperanza matemática de este descenso variable se situaría, con una confianza del 95%, entre el 8’4% y el 39’9%. Para terminar de caracterizar este efecto variable deberíamos hablar también de la varianza del mismo, cuya estimación puntual vale 0’013.

3.- OTRO META-ANALISIS DE ANTIHISTAMÍNICOS.

1. *Odds ratio*.

OR lOR V(lOR) L.I. L.S. WC WCl0R WC(l0RC-l0R)2 WC2 V(ale) WDC WDCl0R

4,13 1,42 1,55 -1,02 3,86 0,65 0,91 0,48 0,42 1,55 0,65 0,91

1,48 0,39 0,07 -0,11 0,89 15,33 5,96 0,40 234,92 0,07 15,33 45,96

1,15 0,14 0,25 -0,83 1,12 4,03 0,57 0,67 16,24 0,25 4,03 0,57

3,10 1,13 0,40 -0,11 2,37 2,52 2,84 0,84 6,33 0,40 2,52 2,84

0,93 -0,07 2,14 -2,93 2,80 0,47 -0,03 0,18 0,22 2,14 0,47 -0,03

0,19 -1,66 1,25 -3,86 0,53 0,80 -1,33 3,91 0,64 1,25 0,80 -1,33

5,45 1,70 0,58 0,20 3,19 1,72 2,93 2,26 2,97 0,58 1,72 2,93

10,00 2,30 1,10 0,25 4,36 0,91 2,09 2,79 0,83 1,10 0,91 2,09

1,81 0,59 0,07 0,08 1,11 14,56 8,63 0,02 212,13 0,07 14,56 8,63

sumas: 40,98 22,58 11,57 474,69 40,98 22,58

Q=11,569 α=0,050 W– =4,553 U=29,398 Ef Fijos D=0,000

P=0,172 Zα/2=1,960 SW2=36,010 (Q-(C-1))/U=0,121 Ef Aleat D=0,121

l0R SE(l0R) LI(l0R) LS(l0R) OR LI(OR) LS(0R)

Efectos Fijos 0,551 0,156 0,245 0,857 1,735 1,277 2,357

Efectos Aleat 0,606 0,228 0,160 1,053 1,833 1,173 2,865

Según el modelo de efectos fijos los antihistamínicos aumentan, respecto a placebo, entre 1’277 y 2’357 veces la disparidad “mejoran >50%” versus “no mejoran > 50%”.

Con el modelo de efectos aleatorios los antihistamínicos aumentan en promedio y respecto a placebo, entre 1’254 y 2’462 veces la disparidad “mejoran >50%” versus “no mejoran > 50%”.

La prueba de homogeneidad no alcanza valores significativos. Parece razonable utilizar el modelo de efectos fijos. Nótese que, el pequeño valor de la heterogeneidad conduce: (1) a resultados no significativos de la prueba; y (2) a intervalos muy similares. Recuérdese que la elección de un modelo u otro es considerado por muchos autores como una premisa, no una consecuencia del análisis.

2.- Riesgo Relativo.

RR IRR V(IRR) L.I. L.S. WC WCIRR WC(IRRC-IRR)2 WC2  V(ale) WDC WDCIRR

3,27 1,19 1,16 -0,92 3,30 0,86 1,02 0,56 0,74 1,18 0,85 1,01

1,28 0,24 0,03 -0,07 0,56 39,10 9,50 0,77 1528,83 0,05 22,22 5,40

1,13 0,12 0,18 -0,71 0,95 5,57 0,67 0,38 31,00 0,20 5,02 0,61

1,86 0,62 0,13 -0,08 1,32 7,76 4,80 0,43 60,21 0,15 6,74 4,17

0,94 -0,06 1,87 -2,75 2,62 0,53 -0,03 0,11 0,29 1,89 0,53 -0,03

0,21 -1,56 1,15 -3,66 0,54 0,87 -1,36 3,28 0,76 1,17 0,86 -1,33

3,33 1,20 0,34 0,06 2,34 2,96 3,56 1,99 8,75 0,36 2,80 3,37

3,08 1,12 0,40 -0,11 2,36 2,51 2,82 1,38 6,31 0,42 2,40 2,69

1,61 0,48 0,05 0,06 0,90 22,02 10,52 0,20 484,66 0,06 15,42 7,37

sumas: 82,18 31,51 9,09 2121,55 56,83 23,25

Q=9,095 α=0,050 W– =9,131 U=56,363 Ef Fijos D= 0,000

P=0,334 Za/2=1,960 SW2=171,396 (Q-(C-1))/U= 0,019 Ef Aleat D= 0,019

lRR SE(lRR) LI(lRR) LS(lRR) RR LI(RR) LS(RR)

EfectosFijos 0,383 0,110 0,167 0,600 1,467 1,182 1,821

EfectosAleat 0,409 0,133 0,149 0,669 1,505 1,161 1,952

Con el riesgo relativo se reproducen los mismos resultados. La interpretación ahora sería: según el modelo de efectos fijos (aleatorios) los antihistamínicos multiplican entre 1’182 (1’161) y 1’821 (1’952) veces la proporción de pacientes que “mejoran >50%” con placebo.

4.- UN M-A SOBRE LA VACUNA DE LA TUBERCULOSIS.

El primer aspecto que resalta al mirar los datos es la diferente prevalencia de la TBC en el grupo control en los diferentes EC (quizás explicable por el diferente año de inicio). Existe una heterogeneidad de salida. ¿Será igualmente eficaz la vacuna en estas situaciones?. ¿Son igualmente aplicables medidas aditivas (riesgo atribuible) que implican descenso lineal de la TBC, y medidas multiplicativas (riesgo relativo, odds ratio) que implican descenso proporcional?.

Al estudiar el m-a basado en el riesgo relativo se observa una marcada heterogeneidad de los resultados debida especialmente al último estudio, que presenta unos resultados de no eficacia (el IC es casi simétrico en el 0) que contrastan con los resultados de los otros seis EC (si bien el segundo y el sexto analizados individualmente no alcanzarían la significación estadística) que apoyan la eficacia. El elevado valor del estadístico Q y su significación estadística confirman esta impresión de heterogeneidad.

La heterogeneidad no varía al cambiar la medida resumen de la eficacia, ya que el m-a basado en el odds ratio reproduce las conclusiones anteriores. Todo ello aconseja utilizar el modelo de efectos aleatorios que apunta a un efecto significativo de la vacuna.

Sin embargo, el estudio 7 tiene unas características (clasificación en TBC más pobre) que lo hacen sospechoso de estar dotado de considerables sesgos de información. Nótese que este estudio (y el sexto) tiene(n) una proporción de TBC extremadamente baja, lo que podría originar su diferente comportamiento, bien por el sesgo citado, o bien porque a estos niveles la vacuna pierde eficacia.

En resumen:

1. BCG es eficaz en las poblaciones representadas por los primeros EC.
2. Se debe aconsejar el estudio de las diferencias que el último EC (y el 6º) tiene(n) con el resto para decidir si, en estas poblaciones, BCG no tiene efecto, o bien no se ha podido establecer por el sesgo de información.

(3) Sería interesante repetir el análisis suprmiendo los dos últimos estudios.

5.- EFECTO DE LA ASPIRINA.

a) Bajo el modelo de efectos fijos,

Q=14’045 < 18’307 = χ210,0’05 [También: P(Q=14’045 >χ210) = 0’173 ]

Nada se opone a aceptar la H0 de homogeneidad del efecto.

[También: no hemos logrado demostrar que existan diferencias entre estudios.]

b) Yln(OR) = -0’284; SEYln(OR)= 0’042

IC95%lnOR = [-0’284±1’96·0’042] = [ -0’367 , -0’201 ]

OR = 0’753 IC95%OR = 0’693 , 0’818

El efecto de las aspirinas, que asumimos único en todos los estudios, consiste en descender entre un 18 y un 31% la razón “presentar/no presentar” estas enfermedades.

(OJO: al ser el OR=0’75, el descenso es del 25%, no del 75%).

c) Ahora asumimos que el efecto varía: Y\* = ln(OR) → N(μlnOR, σ2lnOR)

μlnOR = -0’312 ; σ2lnOR = D = 0’0085

D representa la variabilidad entre estudios del efecto

[No confundir con V(Y\*)=0’00289 que representa el ruido de estimación de la media Y\* del efecto aleatorio Y\*]

d) Esta crítica, válida para la prueba de hipótesis, no es válida para la estimación por intervalo, ya que esta última permite al investigador valorar directamente la magnitud del efecto observado e interpretar si tiene, o no, relevancia práctica.

El nivel de significación informa de la “significación estadística” para rechazar cierta hipótesis, mientras que el intervalo de confianza informa de qué valores del parámetro son plausibles, vistos los resultados observados. Al dar los valores del parámetro, está informando también de la relevancia práctica.

6.- OTRO SOBRE ACIDO ACETILSALICILICO.

a)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | OR | Ln(OR) | V(ln(OR)) | W | W·ln(OR) |
| A | 0’58 | -0’55 | 0’012 | 86’48 | -47’70 |
| B | 1’18 | 0’17 | 0’018 | 52’82 | 8’84 |
|  |  |  | Suma | 139’30 | -38’86 |

Estimaciones: Ln(OR) = -38’86 /139’30 = -0’28

SE(ln(OR)) = √ (1/ 139’30) = 0’085

IC95%ln(OR) = -0’28 ± 1’96 ·0’085 = [-0’45, -0’11]

IC95%OR = exp [-0’45, -0’11] = [0’64 , 0’89] (o bien sus inversas)

b) Que tomar esporádicamente AAS desciende entre un 11% y un 36% el riesgo.

Que el efecto estimado es único, constante. La amplitud refleja incerteza, no variabilidad del efecto.

[NOTA: esto no tiene demasiado sentido viendo que los IC individuales no se solapan o que la prueba de homogeneidad [Q= 165’67 > 3’84) es altamente significativa]

c) El intervalo de confianza del OR incluye el valor 1 de independencia, por lo que no han encontrado diferencias significativas.

d) Existen demasiadas diferencias entre A y B para creernos que es de efectos fijos:

tipo de estudio experimental (EC) u observacional (cohortes)

hombres o mujeres

dosis f ijas o dosis muy variables....

7 – INFECCIÓN POR HELICOBACTER PYLORI Y DISPEPSIA FUNCIONAL.

METANÁLISIS DE LA EFICACIA DEL TRATAMIENTO ERRADICADOR

W¯ = 90’69/9 = 10’077

V(Y) = 1/90’69 = 0’011 SE(Y) = **0’105**

Y = 13’29/90’69 = **0’147** OR = exp(0’147) = **1’16**

IC(Y) = 0’147 ± 1’96·0’105 = **[-0’0588, 0’3528]**

IC(OR) = exp[-0’0588, 0’3528] = **[-0’9429, 1’4231]**

Q = **13’29** V(W) = 1165’38–9·10’0772)/8 = **1051’14**

U = 8·(210’077-1051’14/90’69) = **1587’8922**

(Q>C-1) D = (13’29-8)/1587’8922=**0’00333**

La homogeneidad es dudosa, P ≈ 0’01 (la variancia estimada es mínima).

Si la incorporamos para un modelo de efectos aleatorios:

(los pesos se modifican: W\*c = (0’00333 + Wc-1)-1)

Y\* = 8’91/49’76 = **0’179** OR = exp(0’179) = **1’20**

V(Y\*) = 1/49’76 = 0’0201 SE(Y\*) = **0’1418**

IC(Y\*) = 0’179 ± 1’96·0’1418 = **[-0’099, 0’457]**

IC(OR\*) = exp[-0’099, 0’457] = **[0’906, 1’579]**

8.- PARA PENSAR

Por supuesto, deben incluirse todos los EC, ya que seleccionarlos en función de los resultados introduce sesgos. Nótese que la tendencia de los autores y de las revistas científicas a publicar sólo los resultados positivos –en el sentido de estadísticamente significativos- hace que las bases bibliográficas no sean una fuente suficiente para el meta-análisis.

El m-a de efectos fijos establece que el efecto es el mismo en las diferentes condiciones de administración, por lo que será más fácil conocer cuál es el efecto esperado en cada caso.

Para que un estudio aislado permita inferencia causal debe recurrir a las premisas formales estadísticas: suficiencia del modelo, asignación aleatoria, ausencia de variables confusoras o asignación ignorante del tratamiento.

Un conjunto de información sobre un mismo tema, obtenida de diversas fuentes, permite recurrir a los factores de Hill: Temporalidad, seudo-experimentación, gradiente biológico, intensidad de la relación, especificidad, consistencia, plausibilidad, coherencia, analogía,....

9.- Meta-Analysis of Total Coronary Heart Disease Events

La prueba de chi-cuadrado para la hetereogeneidad del efecto es claramente significativa, X2=18,07, con un p-valor del 0,0089, lo que explica la etiqueta de “random” del gráfico (se ha asumido un modelo de efectos aleatorios).

Todos los ensayos ofrecen estimaciones puntuales a favor de aspirina. La combinación de los 5 ensayos da a lugar una estimación conjunta que se lee cómo que la Aspirina es un factor protector sobre los eventos cardiovasculares, con un OR del 0,72 e IC95% de 0,60 a 0,89.